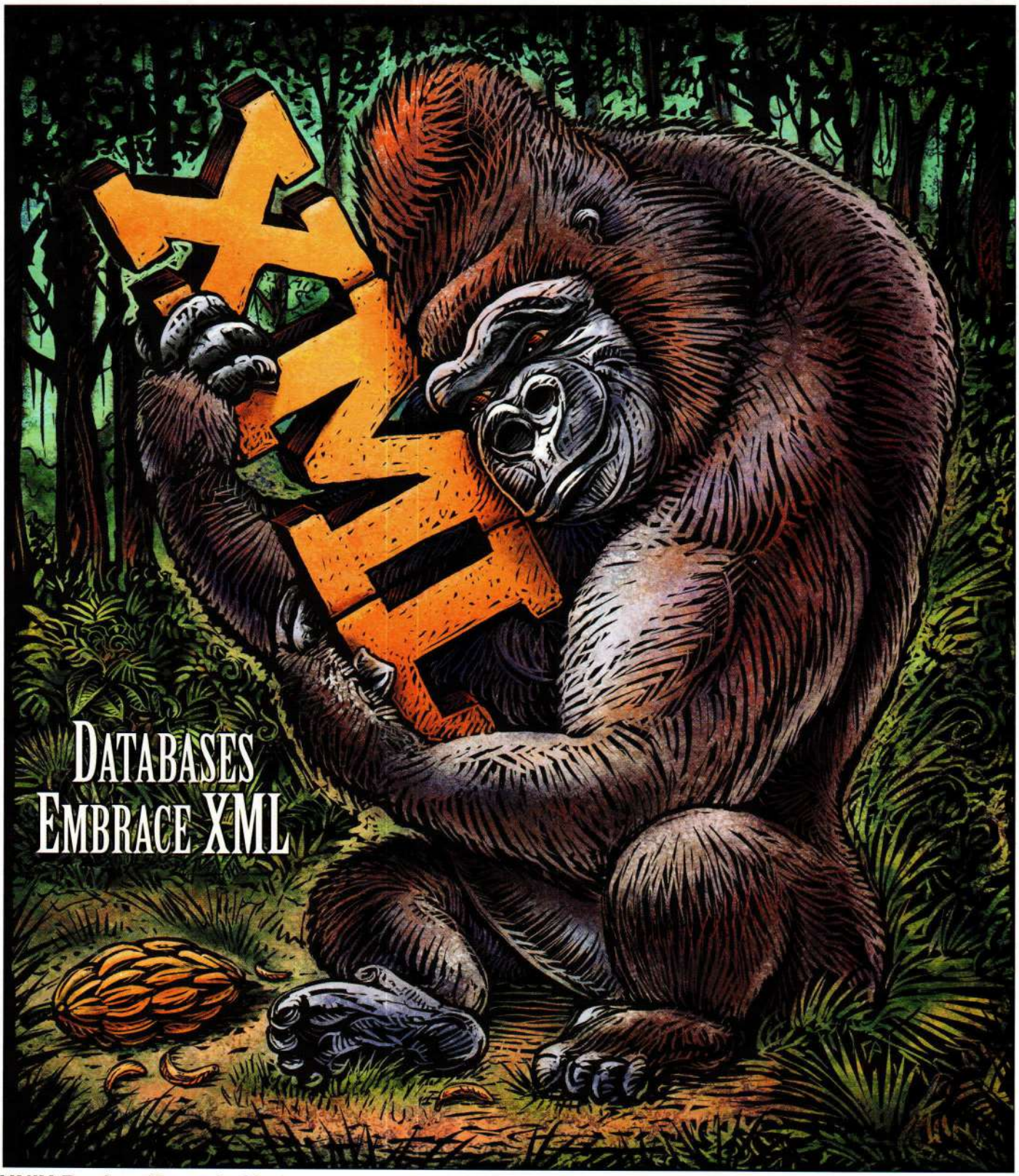
Server/Workstation EXPERT

AUGUST 2001
Vol. 12 No. 8 $5.50
http://swexpert.com

Server-Based Computing Solutions



UNIX Basics: Domain Names

Work: The Nesting Instinct

Fibre Channel Technologies Conference 2001

Leading the Way in High-Speed Technology

REGISTER ONLINE NOW

Storage Area Networks ○ Storage over Ethernet and IP ○ RAID Systems ○ Storage Management Software ○ SAN Management ○ SAN Applications ○ High-availability Systems ○ High-speed Versions ○ Global SANs ○

www.fctexchange.com

SAN/LAN/MAN/ WAN connections ○ Managed Storage Providers ○ Network-attached Storage ○ Security ○ Clustering ○ Cabling/Interconnect Issues ○ Testing ○ Marketing Analysis ○ Infiniband

Conference: October 16-18, 2001
Exhibits: October 16-17, 2001
Tutorial: October 15, 2001
San Jose, California
Doubletree Hotel

Association Sponsor: FCIA

fibre channel technologies conference & expo

www.fctexchange.com

Register Online and SAVE $100!
Register before September 17 and SAVE $200!

For Exhibit Sales and Sponsorship Opportunities:
Contact - Anna Novak
+1.972.714.0514 x2348
a.novak@fctexchange.com

We'll Give You Control Over The Center Of *Your* Universe!

Sun Certified SOLARIS READY

High Performance "Fault Tolerant" Alternatives to Terminal Servers

ConsoleServer 3200 - Delivers centralized management of serial command console ports through direct network access, multiple terminals, and via optional modem. The ConsoleServer 3200 is configurable to 32 devices/17 users and offers features that terminal servers simply cannot provide such as remote dial-up and telnet access, 64k per port audit trail, and login ID/Password protection forn unauthorized access.

ConsoleServer 800 - Same high performance features as the ConsoleServer 3200 in a smaller factory configured 8 user format with 8 device ports, one terminal connection, one network connection and an optional modem port.

A KVM Switch With Stellar Performance

ServerSwitch 8200 - Active, microprocessor controlled switching of keyboard, video and mouse signals among two users and eight Sun computers. Signals are switched transparently to each computer, eliminating interruption of active server processors. The ServerSwitch 8200 is fully compatible with Sun keyboards, mice and graphics monitors, has dedicated cascade ports and comes with both 13W3 and HD15 connections for easy installation with standard Sun cables.

The Best Solutions Under The Sun. Call Us At : 800-871-9838

Lightwave Communications
www.lightwavecom.com
100 Washington Street, Milford, CT 06460

Europe:
E-mail: office@lightwave.de
Phone: +49-89-306-3810
Australia: +61-39-652-7226

Server/Workstation EXPERT

The leading publication for server-based computing solutions

Contents

PAGE 8

8 News

Linux on IBM eServers Complete; HP Builds Java Framework; Sun Broadens Java; Itanium Is Here; SAN Vendors Try Cooperation; IBM Shark and Hitachi Freedom Cozy Up.

PAGE 42

42 Cover Story

by Alexandra Barrett, Senior Editor

Databases Embrace XML

Whether as a data transport or native storage scheme, databases, often called the 900-pound gorilla of IT, are making room for XML.

PAGE 48

48 Product Review

by Ian Westmacott, Technical Editor

Traffic Analysis on the Threes

Three vendors vie to be your "eye in the sky" for Web site usage.

Departments

6 Editorial
56 New Products
61 Server/Workstation Marketplace
72 Advertisers' Index

BONUS DISTRIBUTION

For Advertisers of the **September** Issue:

Internet World Fall
New York, NY October 3-5

COMING NEXT ISSUE

Datawarehousing for the Enterprise

SERVER/WORKSTATION EXPERT Magazine (ISSN# 1524-4849) is published monthly by Computer Publishing Group, 1340 Centre St., Newton Center, MA 02459. Telephone (617) 641-9101. Periodicals Postage Rates paid at Boston, MA, and at additional mailing offices. Posted under Canadian IPM #0235873. This publication is free to qualified subscribers as determined by the publisher. Subscription rates are $60 per year in the United States, and $75 (surface mail) and $125 (air mail) outside the United States. Subscription requests can be sent to Circulation Department, *SERVER/WORKSTATION EXPERT* Magazine, 1340 Centre St., Newton Center, MA 02459, or electronically mailed to circ@cpg.com.

POSTMASTER, please send all address changes to *SERVER/WORKSTATION EXPERT* Magazine, Circulation Department, 1340 Centre St., Newton Center, MA 02459. Please allow 6-8 weeks for change of address. Include your old address as well as new–enclosing, if possible, an address label from a recent issue.

All rights reserved. © Copyright 2001, Computer Publishing Group. No part of this publication may be transmitted or reproduced in any form by any means without permission in writing from the publisher.

Material for publication should be sent to the attention of Doug Pryor at the above address or electronically mailed to dpryor@cpg.com. Letters sent to the publication become the property of the publication and are assumed to be intended for publication and may be used so. All information herein is believed to be accurate to the best of our ability.

Columns

14 Ask Mr. Protocol

by Michael O'Brien

Mr. Protocol and Linux to Go

Mr. P. discusses the Linux PDA movement–and then some.

20 UNIX Basics

by Peter Collinson

Domain Names

Our UNIX guru provides an in-depth look at the Domain Name game.

24 Storage

by Henry Newman

Secondary Storage

Our storage maven discusses different storage media, including helical and linear tape, as well as optical disks.

28 Work

by Jeffreys Copeland and Haemer

The Nesting Instinct

The Jeffs write a yacc parser for XML while switching from vi to emacs.

32 Java Class

by Jim Frost

Postal: Design Considerations

This month in Postal, Jim helps improve reliability and forms the foundation for later performance and feature enhancements in the IMAP e-mail client in Java.

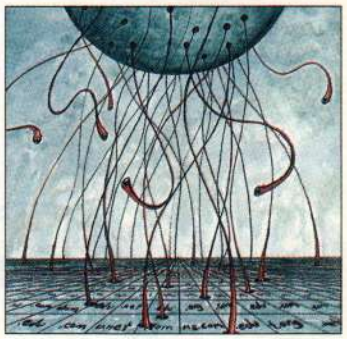
38 Open Sources

by Cameron Laird

Fun with SOAP

Cameron gets his hands dirty by demonstrating some examples of SOAP.

PAGE 14



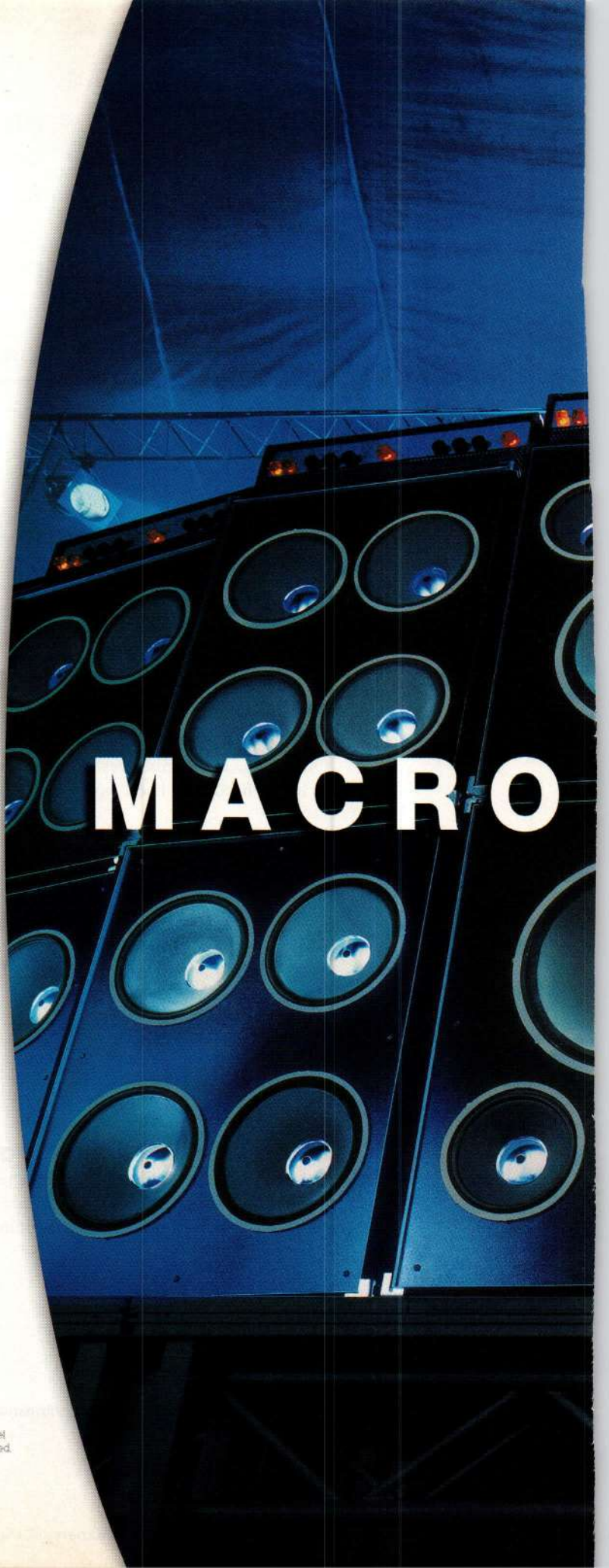
PAGE 20

PAGE 28

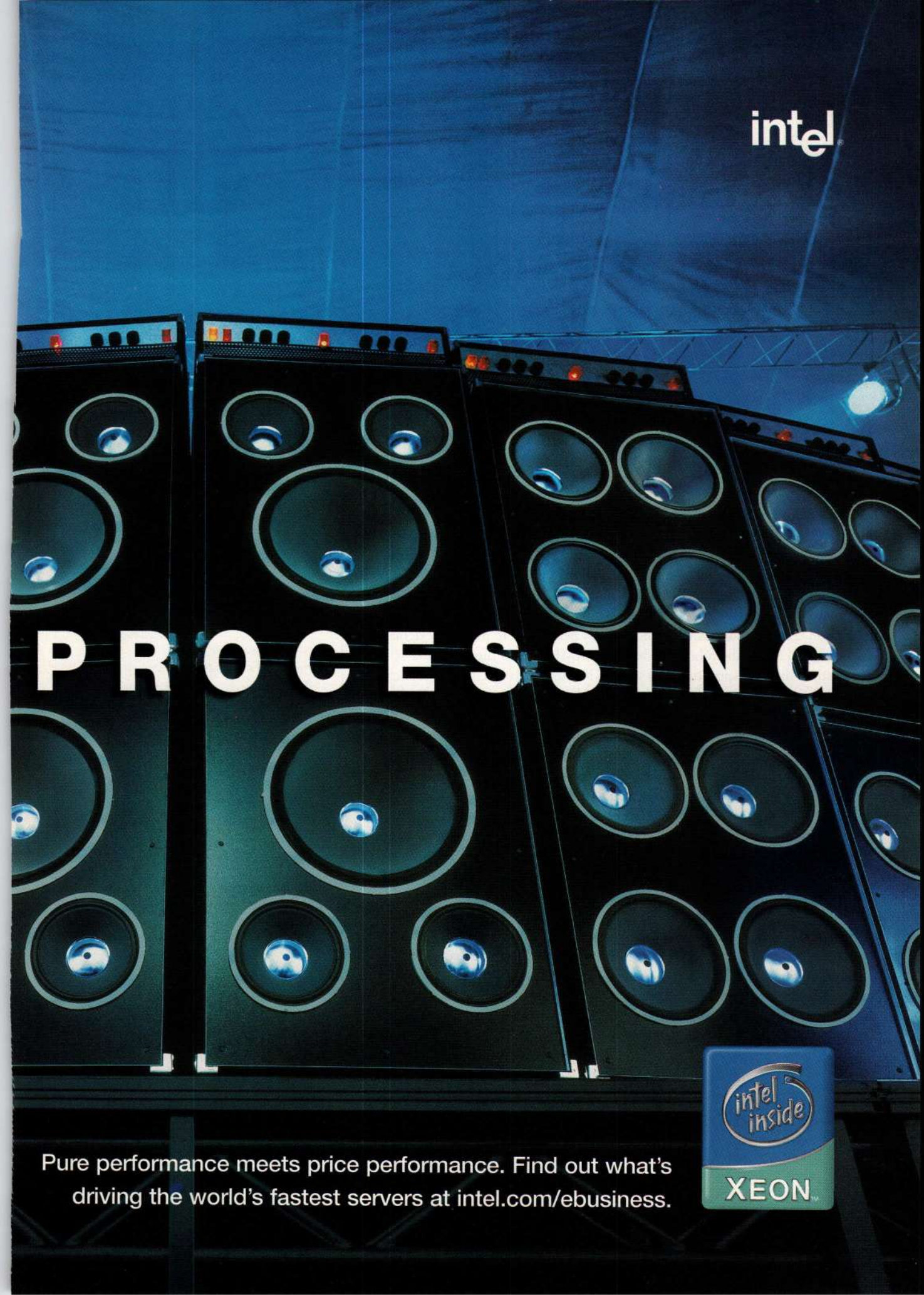
PAGE 32

PAGE 38

MICROPROCESSING

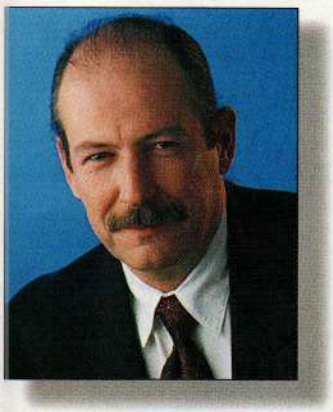


© 2001 Intel Corporation. Intel and Xeon are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. All rights reserved.


intel
PROCESSING
Pure performance meets price performance. Find out what's driving the world's fastest servers at intel.com/ebusiness.
intel inside
XEON

EDITORIAL

dpryor@cpg.com



A Unique 3-Year-Old

Has it been just three years since the World Wide Web Consortium (W3C) recommended version 1.0 of the eXtensible Markup Language? You wouldn't know it from all the work that's going on with this fledgling scheme–if you can call it that. It seems to be one of those rare birds that not only gets a lot of attention but is also actually worthwhile.

This month's cover story "Databases Embrace XML," Page 42, by Alexandra Barrett, explores how a 900-pound gorilla app is being changed by XML. As Alex says, "If the early days of XML were marked by organizations using it as a data interchange mechanism–for a business-to-business or enterprise application integration context–today, XML is increasingly being used to mark up original content." And there's the rub. How deep does XML go into the database? Is it a convenient data transport or a new scheme for storing data? Does XML mean a new data infrastructure or can existing databases be adapted to XML data?

There is no doubt that the body of XML data is growing. According to Robert Perry, senior analyst of the Yankee Group, "XML is the dominant way of describing content for Web applications. It's already there, not just moving in that direction." The key issue Alex addresses is "how does one manage all these XML documents?" To find out, you'll just have to read her story.

If you have experiences with XML and databases that you would like to relate, we would love to hear from you–especially if you are debating natively storing XML within a relational database environment. After all, no story is the final word on a subject. E-mail dpryor@cpg.com.

Also in this issue, Ian Westmacott does duty as a traffic analyzer. In "Traffic Analysis on the Threes," Page 48, he puts three Web log analysis packages through some rigorous tests. They include Urchin 3 from Quantified Systems, NetTracker 5.0 from Sane Solutions, and Enterprise Reporting Server 3.5 from WebTrends. For each he looks at the salient features, discusses how to use them, and explains how each reports results. If you're looking for a reporting tool, you should pay particular attention to Ian's "Notes on Benchmarking," Page 54.

Doug Pryor

SW EXPERT

August 2001 Vol. 12 No. 8

Publisher S. Henry Sacks
shs@cpg.com

Editor-in-Chief Douglas Pryor
dpryor@cpg.com

Managing Editor John Chilson
jchilson@cpg.com

Technical Editor Ian Westmacott
ianw@cpg.com

Senior Editor Alexandra Barrett
alex@cpg.com

Production Editor Domenic Argentieri
domenic@cpg.com

Art/Production Director Brad Dillman
bdillman@cpg.com

Production Assistant Chrystie Hopkins
chrystie@cpg.com

Circulation Manager Tina Jackson
jamal@cpg.com

Accounting Manager Robin Yearwood
robin@cpg.com

Administrative Assistant Karen Parrell
karen@cpg.com

Business Manager Susan R. Sacks
srs@cpg.com

Editorial Offices
Computer Publishing Group
1340 Centre Street
Newton Center, MA 02459
(617) 641-9101

World Wide Web
http://www.cpg.com

COMPUTER PUBLISHING GROUP

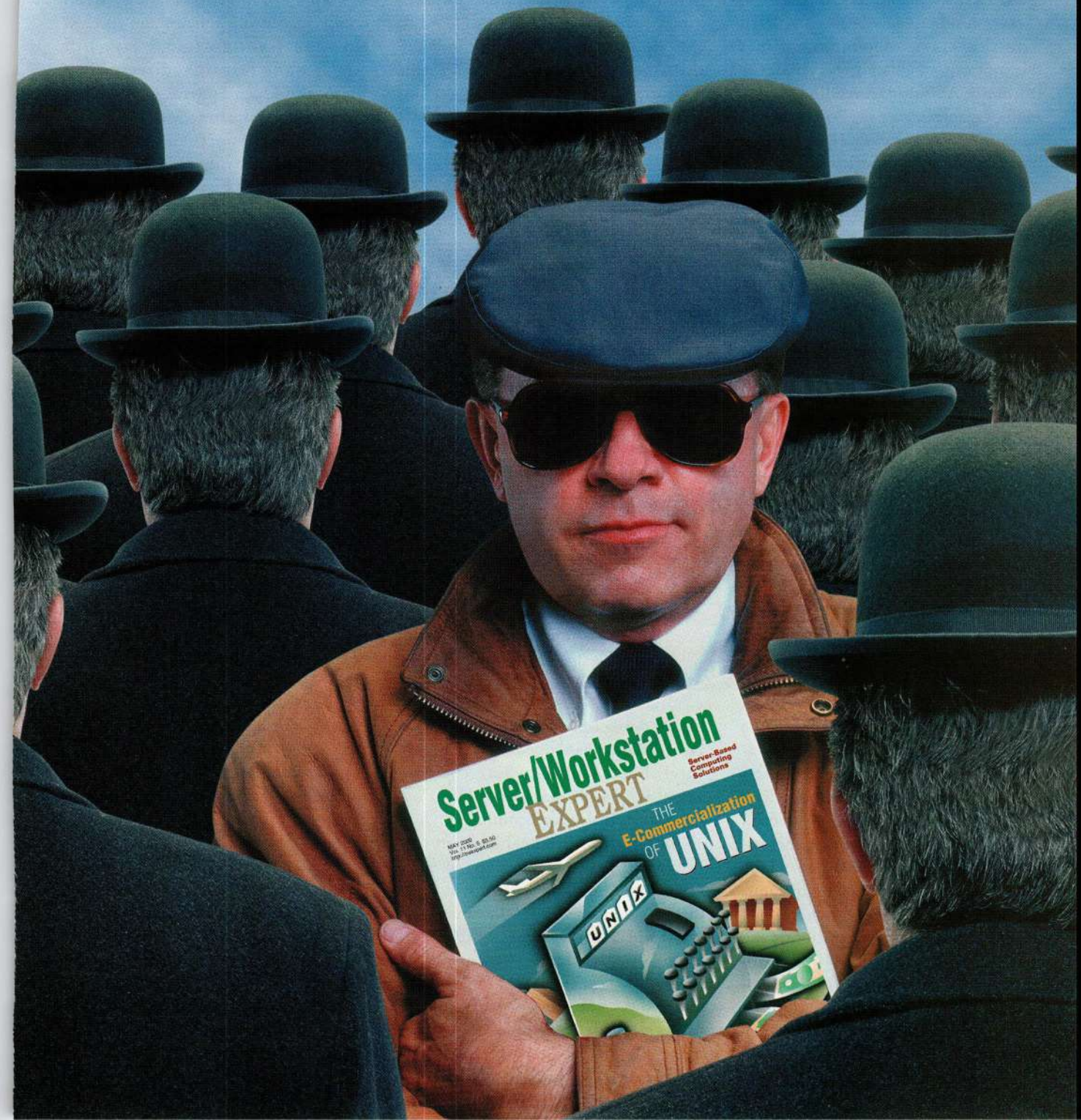
BPA INTERNATIONAL

Printed in USA

One of a Kind

The unique magazine for IT professionals

Don't miss a single issue of *SW Expert*. Simply fill out the subscription card and drop it in the mail.



Linux on IBM eServers Complete

The plan IBM Corp., Armonk, NY, embarked on last year to Linux-enable its entire line of eServer products has been completed with this May's announcement that the company's high-end eServer iSeries (formerly AS/400) server line is now Linux-capable.

Linux also runs on IBM's Intel-based xSeries, the pSeries (formerly RS/6000), and zSeries S/390 mainframes, which the company opened up to developers this spring through its Linux Community Development System initiative. The free program gives developers access to a "virtual" mainframe over the Internet, as well as access to an IBM Shark storage system.

iSeries customers will be able to partition their systems into 31 separate instances of Linux, says Dan Frye, director of the IBM Linux technology center. iSeries users will be able to consolidate applications onto a single system.

The iSeries is typically surrounded by a bevy of application servers, Frye explains. "Now that Linux runs on the iSeries, you can collapse that entire [application] tier onto a single partitioned server." The result is a system that is easier to manage and cheaper to run, Frye says.

With Linux support, iSeries users will also benefit from the wealth of applications currently being written for Linux, which Frye believes is increasingly the default platform to which developers write. "The beauty of Linux is that it's the same everywhere, and that you can get it for all platforms and from all vendors." This June, IBM announced a 30% increase in the number of applications running internally on their eServer line, bringing the total number up to 2,400.

Supporting Linux on the iSeries was largely a matter of porting the operating system to a new chipset. Linux already runs on IBM's POWER chip (formerly PowerPC), but only on the 32-bit version powering the pSeries, and not the iSeries' 64-bit version. Working out issues related to the 64-bit POWER chip was "the key bit" of doing this port, says Frye. "The PowerPC just hasn't been as utilized as Intel, but 64-bit is where the real action is." 64-bit support for Linux on the pSeries is also forthcoming.

STEPHEN SCHILDBACH

Of course, running Linux on the iSeries isn't for everyone. Users in need of a mature operating system will want to continue using OS/400 as their primary operating system, which Frye claims surpasses Linux's capabilities in terms of workload management, availability, and systems management, to name a few. "Linux is clearly not the equal of OS/400," he says. Users can run both at the same time, in different partitions, though.

Still, analysts are optimistic that customers will take advantage of Linux for the iSeries. Linux on the zSeries System/390 mainframe has been wildly successful, "much more successful than I anticipated," says Bill Claybrook, research director for Linux and Open Systems at Aberdeen Group, Inc., Boston, MA. Not only are zSeries customers running Linux on existing zSeries installations, but IBM has even sold mainframes designed from the get-go to run Linux, Claybrook says.

But it is doubtful that Linux's success on the zSeries will be replicated to the iSeries, says Claybrook. Thanks to the zSeries virtual machine capability, the mainframe can be partitioned into a virtually limitless number of separate instances of Linux. TCO (total cost of ownership) studies have shown that for large installations–service providers, for example–running Linux on a mainframe is more cost effective than running it on several Intel-based servers, Claybrook says. The iSeries, on the other hand, can only be logically partitioned into 31 instances of Linux, reducing its appeal for large scale ISPs and ASPs. "I really doubt we'll see the same groundswell of support for Linux on iSeries that we saw on System 390."–*ab*

Sun Broadens Java

Sun Microsystems Inc., Palo Alto, CA, sharpened its chief weapon, Java, in its ongoing tussle with Microsoft Corp., Redmond, WA, for the attention of third party developers. At the JavaOne conference held in San Francisco in June, the supplier unveiled

new enhancements, including support for XML in the Java 2 Standard Edition (J2SE) platform, a reference implementation of the Web-based Enterprise Management (WBEM) standard, and a new Peer-to-Peer (P2P) initiative.

Improvements were designed to drive the technology's acceptance into more markets by making it simpler for Java programmers to take advantage of emerging application development technologies.

The latest release of Java Standard Edition 1.4 adds XML support to the core platform. XML has become a key means of easing problems that arise when applications exchange information because it is often in incompatible formats.

The new JAX Pack is a collection of XML-based Java APIs:

- Java API for XML Processing (JAXP) provides interfaces that parse and process XML documents, and a transformation engine that works with XML Style Sheets.
- Java API for XML Registries (JAXR) defines ways for Java applications to access XML registries, such as ebXML and Universal Description, Discovery, and Integration (UDDI).
- Java API for XML Messaging (JAXM) provides Java support for sending and receiving Simple Object Access Protocol (SOAP) messages.
- Java Architecture for XML Binding (JAXB) helps to map XML data into Java objects.
- Java API for Remote Procedure Call (JAX-RPC) supports the use of XML for remote procedure calls, where a procedure call on one computer is transmitted over a network as XML and then delivered to another computer.

XML has become a key means of easing problems that arise when applications exchange information.

"The JAX Pack is an important addition because it forms the missing link between Java Web programming and the XML data transforms needed for application interaction to create Web services," says Martin Marshall, managing director, at Zona Research Inc., a Redwood City, CA, market research firm.

The new interfaces should cut program development time. "If a programmer was working with a simple document, he may have to write 300

IBM Shark and Hitachi Freedom Cozy Up

Customers looking for a suitable high-end storage solution for their UNIX and mainframe systems should consider a recent announcement from IBM Corp., Armonk, NY, and Hitachi Data Systems Corp. (HDS), Santa Clara, CA. This June, the two companies announced they will cross license key APIs for their storage systems, IBM's Enterprise Storage Server, a.k.a. Shark, and Hitachi's Freedom Storage Lightening 9900 series.

Specifically, IBM will license to Hitachi key APIs to its Peer-to-Peer Remote Copy (PPRC) and Extended Remote Copy (XRP) functions, as well as allow Hitachi to incorporate IBM's FlashCopy, Multiple Allegiance (MA), and Parallel Access Volumes (PAV) functions into the Freedom storage subsystem.

Hitachi, in turn, will license APIs for the Lightening 9900 to IBM, allowing IBM to support functions such as Hitachi's NanoCopy and ShadowImage on Shark. Hitachi has also indicated that it will retrofit older storage products with the IBM functionality.

Cross licensing these APIs will allow users of one system to purchase storage from the other company, while maintaining application-level compatibility between the two. Interoperability between storage systems is a "major hot button for customers," says Mike Harrison, IBM director of business alliances. The IBM/Hitachi announcement allows customers "to put two systems side by side, link them up, or replace them with one another, while still using the same software."

This level of interoperability should appeal to buyers of enterprise-class storage servers–especially in shops using both UNIX and System 390–says John Young, research program director for enterprise storage at analyst firm D.H. Brown, Port Chester, NY. "Virtually all customers want to have more than one vendor in their shop," he says. By choosing storage from IBM or Hitachi, therefore, customers insulate themselves from becoming tied to a single storage vendor.

From the end-user perspective it appears that IBM and Hitachi's cross licensing announcement is a surefire win. The question remains, however: what's in it for the vendors?

From Hitachi's perspective, the benefit of this strategy is clear: access to previously inaccessible IBM S/390 mainframe functions, says Young, notably PAV and MA, functions implemented on the mainframe to enhance system I/O.

For IBM, however, the benefits are less obvious. Admittedly, says IBM's Harrison, "this does offer the opportunity to increase competition" between IBM and Hitachi, by eliminating the competitive advantage inherent in having sole product compatible with a customers' existing systems.

What the cross-licensing really does, though, is "put added pressure on EMC," says Harrison, whose Symmetrix supports IBM's PAV and MA, thanks to the reverse-engineering skills of its employees, not a licensing arrangement from IBM.

And pressure EMC, these companies must. According to recent data published by Dataquest, EMC Corp., Hopkinton, MA, ended 2000 with a full 34.6% of the external RAID market–exceeding the combined shares of Compaq Computer Corp., IBM, Hewlett-Packard Co. and Sun Microsystems Inc., its closest competitors.

Harrison adds that IBM would happily work with EMC to provide a similar cross licensing deal. "Not only would we consider it, we would welcome it with open arms," says Harrison. "But frankly, I don't expect them to come to the table."

In the meantime, while IBM and Hitachi wait for EMC to open up its Symmetrix systems to the competition, IBM and Hitachi can present customers with a compelling argument for their systems over EMC's, says D.H. Brown's Young: insured interoperability over the long term. "It's a subtle strategy," he says, that ultimately allows them to "lower the hammer over the head of the No. 1 guy."–ab

lines of code to format it from Java to XML," says Anne Thomas Manes, director of marketing innovation at Sun. "Now, that task would require only one line of code."

The Distributed Management Task Force (DMTF), a Portland, OR, consortium, has been the force behind WBEM, which is designed to provide a common framework for building system and network management applications. As a founding member of WBEMSource, which is an open-source implementation of a Java version of the WBEM specification, Sun has been working with other companies to turn the specifications into open-source code that any third party can enhance.

Sun's P2P initiative, dubbed JXTA (pronounced "juxta"), consists of a set of protocols that let computer nodes find each other, organize into peer groups, exchange messages within groups, and manage the nodes forming the group. The JXTA core is a set of programs that use the protocols for these purposes.

One program is peer monitoring, used to manage the nodes in a group. This program detects if one node is using too much bandwidth or capacity, then passes this alert to a management console so an administrator can reconfigure the faulty node.

Running above such core services are generic programs that make use of the underlying protocols. One example is JXTA Search, which consists of source code that performs a distributed search among JXTA nodes. Above the services run applications, such as instant messaging and a groupware program, that would use the lower levels of JXTA to let peers have these functions.

Sun has made its JXTA services available through the open-source model, but will face competition. In February, a group of vendors, (Intel, Entropia Inc., and Engenia Inc.) formed the Peer-to-Peer Working Group (P2PWG), an ad hoc consortium to develop P2P standards. "We felt the P2PWG was not well organized–it seemed like Intel ran the group–so we decided not to participate in it for at least the moment," says Sun's Thomas Manes.–*paul korzeniowski, freelance writer*

HP Builds Java Framework

Hewlett-Packard Co., Palo Alto, CA, is helping to ease Java application development by releasing a new application development framework and Web server. The company's plans, which were unveiled at the JavaOne conference in San Francisco in June, are to make it simpler for third parties to deliver applications by paring down the wide list of potential Application Programming Interfaces (APIs) to a more manageable number.

STEPHEN SCHILDBACH

"There is a wide variety of services, like messaging, that every Java application needs," says Daryl Plummer, group vice president at Gartner Group Inc., Stamford, CT. "HP has made it easier for developers to access and use those services by delivering a set of common interfaces to them."

As corporations try to build new applications that exploit Internet technology, Web services have been getting a lot of attention. Sun Microsystems Inc., Palo Alto, CA, and Microsoft Corp., Redmond, WA, have emerged as the top two contenders in that space by delivering a wide range of tools and standards to ease development. HP's work may seem similar, but it is not: the company has delved deeper into application infrastructure, focusing on capabilities programmers use to move information from servers to browsers, according to Plummer.

The HP Core Services Framework (CSF) defines how a variety of services (messaging, security, naming, directory, logging) interoperate with each other and with non-Web based systems. The framework outlines practices for code sharing and reuse, streamlining development, and a uniform approach to life cycle processes, management, and configuration of application services.

In addition, CSF removes the complexities of developing service-based applications by providing a core set of APIs to common services. Currently, developers select from different options and then various components–a difficult process because new APIs quickly evolve and appear almost weekly. The HP framework pares the potential number to a select few and provides programmers with a common means of invoking these services, according to HP.

Since it's a Java-based framework, CSF supports a variety of Java standards and services including Java Management Extension (JMX), Java Naming and Directory Interface (JNDI), and Java API for XML Parsing (JAXP). These interfaces let users connect to authentication services (such as Lightweight Directory Access Protocol), databases, and XML files.

Developers and independent software vendors (ISVs) can now download, at no charge, and use HP CSF as a foundation for building Web-enabled application services. CSF code can be embedded within DBMSs, applications, servers, devices, appliances, and other applications. In addition, the framework supports the run-time configuration and management of applications.

The framework is one of the first reference implementations of the Java Services Framework as described in Java

Ultimate Performance, Infinite Possibilities.

Introducing Rave Systems Redundant RackMount-4UAX2200 featuring UltraSPARC-III technology

Rave Systems Redundant RackMount-4UAX2200

(shown with face plate removed)

Original Sun Netra™ AX2200 with single or dual 750MHz UltraSPARC-III processors.

This powerful machine is designed to provide exceptional power for any application that requires reliable and dynamic compute intensive performance. Rave Systems Redundant Rackmount 4UAX2200 features:

- Redundant auto-ranging power supplies
- Up to two 750MHz UltraSPARC-III processors
- Up to 8GB of RAM
- Four 32 or 64-bit PCI slots
- Flexible disk, tape and removable media options
- Host of standard interfaces including Fibre, Firewire, Ethernet, UPA, Serial and Parallel ports

Call Rave Computer today at 1-800-966-7283 or go to www.rave.com.

RackMount-1UAX1105

Entry level reliable UltraSPARC-IIe **server** in a 1U form factor. Featuring **Netra™** AX1105 with two front **accessible IDE or SCSI disk drives, one PCI** slot and **up to 2GB of RAM.**

Redundant RackMount-2UAXi-C

Featuring Ultra™ AXi, a robust and reliable UltraSPARC-IIi server in a 2U rack with four removable hard drives and dual hot-swap power supplies. Supports high performance graphics, and has integrated dual-channel Ultra-Wide SCSI.

Redundant RackMount-4UAXmp

Featuring Ultra™ AXmp motherboard that provides raw **multi**-processing UltraSPARC-II power in a redundant 4U **rack.** Supports up to 4GB of RAM and has six PCI slots.

Redundant Disk Array

Fault resilient disk array with multiple bus configuration options and RAID capable.

ULTRASPARC DRIVEN

With Rave Systems, you get the best of both worlds - engine by Sun and solutions by Rave Computer. You get the power and reliability of Sun, and from Rave Computer, you get quality flexible integration solutions with service. Call us today and tell us what you need at 1-810-939-8230 or send an e-mail to info@rave.com.

SOLARIS

Rave Computer Association, Inc.

1-800-966-7283

More information about Rave Computer Association and our Rave Systems can be found at: www.rave.net

Fax: (810) 939-7431 E-mail: info@rave.com

Leasing Available. Call 1-800-500-7283

Sun AUTHORIZED VALUE ADDED INTEGRATOR

ISO 9002

Sun, Sun Microsystems, and the Sun logo are trademark or registered trademarks of Sun Microsystems, Inc., in the United States or other countries. All other trademarks are the property of their respective companies. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the United States and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc. All other products or service names mentioned herein are trademarks of their respective owners.

Specification Request 111, a specification developed by the Java Community Process (JCP), an ad hoc group which develops and promotes Java standards. HP says CSF will continue to support new Java APIs as they are added to the JSR 111 specification.

Consequently, Gartner Group's Plummer thinks the HP framework will be successful. "Currently, no other vendor has tried to address the programming issues that HP framework is tackling," he says. "Also, having JCP approval is important because it makes the framework an industry standard and not a vendor's proprietary implementation."

The company's first use of CSF is in the HP Internet Server. This server is targeted at ISVs and developers building server applications and Web services requiring HTTP communications, as well as execution of Java Server Pages (JSPs) and Java Servlet code.

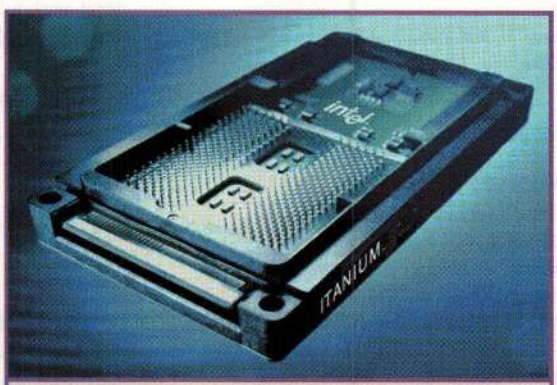
The HP Internet Server implements the JSP 1.2 and Java Servlet 2.3 standards. JSPs separate Web page design from business logic, letting designers change page layout without affecting the underlying dynamic content. As server-side applets, servlets provide developers with the ability to extend the functionality of a Web server and provide access to existing back-office data sources.

"We think the HP Internet Server meets the need of companies that want a high-end Web server application or a low-end application server," says John H. Capobianco, general manager, marketing strategy, HP's middleware division. The HP Internet Server is expected to be available for free downloading in mid-July.–*paul korzeniowski, freelance writer*

Itanium Is Here

With the launch of the 64-bit Itanium processor, Santa Clara, CA-based Intel Corp. might have just christened a new era of computer processing. Itanium, which is designed to handle processor intensive applications such as e-commerce encryption and decryption, graphics, and scientific modeling, will be shipped as part of new systems from an expected 30 computer manufacturers including IBM Corp., Dell Computer Corp., Compaq Computer Corp., as well as Hewlett-Packard Co., which worked with Intel to create the new computing architecture known as EPIC (Explicitly Parallel Instruction Computing), which Itanium uses. EPIC is said to enhance Itanium's ability to process parallel data streams.

But, the pundits say, we won't know for sure if a new era of computer architectures is dawning for some time. According to Rich Partridge, analyst at D.H. Brown, Port Chester, NY, the first systems to hit the market are all very much the same. "What the vendors use as their rational is you don't want to have a lot of changes [at this stage]. Nearly everyone is taking the Intel reference platform that Intel has designed as a proof point and have picked that up as the foundation. The kind of differentiations you'll see probably next generation isn't there right now."



Intel's 64-bit Itanium processor.

According to Partridge, part of the reason for the similarity across the first generation of products is system vendors don't want to worry about fighting bugs with the hardware at this point. He also notes it is important to remain somewhat standard and uniform so users can get their code ported to 64-bit addressing. "There are a lot more fundamental rudimentary basic instructions that the EPIC compiler needs to compile to have the hardware execute multiple things simultaneously," says Partridge.

One issue early users face is the operating system choice. For example, at least four operating systems will support Itanium-based systems, including 64-bit versions of Microsoft Windows, HP's HP-UX 11i v1.5, IBM AIX-5L, and Linux. But Microsoft hasn't released a full beta version of its 64-bit operating system. The initial shipment of Windows 2000 64-bit Limited Edition is only supposed to run on an 8-way server. So if you want to run a 16-way server on a single system image, making use of all the resources on one image, you really need to run it in UNIX right at the moment, says Jean Bozman, research director, IDC Worldwide Server Group based in Mountain View, CA.

Bozman adds, "If you don't compile your software it is not going to run well on Itanium. You have to run through a compiler for a different hardware target. The chip is a 64-bit microprocessor. It does require a 64-bit OS. And the vendors can quibble with this, but you should be running 64-bit applications. It's pretty much an all 64-bit environment from what I can see."

Some of the first systems to roll out include Dell Computer, which plans to make its PowerEdge 7150 available later this year. IBM's Itanium offering is the eServer x380 available with up to four processors and designed for data mining and data warehousing, online transaction processing and security applications. IBM is also offering the IntelliStation Z Pro workstation. HP has introduced the RX 9610 server, a 16-way server, and the RX-4610, which comes with two to four Itanium processors. Sun Microsystems Inc. does not plan to develop Itanium based systems. "What Sun has said, [is] they want to focus on having everything out of one operating system environment and one chip architecture," says Partridge. "Sun doesn't want to dilute themselves by trying to serve too many different areas." But as early as October 1999, Sun announced that it had the Solaris Operating Environment running on engineering prototype systems based on Intel's Itanium processor.

While Itanium is just starting to hit the market, the processor is said to offer many benefits. One strength is in transaction processing which is important in e-commerce and handling SSL requests. "They have found that Itanium does encrypting and decrypting very quickly," Bozman says. Itanium systems are also designed to serve the scientific space well with floating-point and large data sets for compute-intensive applications. In the area of scalability, Itanium systems can scale up to 512 processors.

The collaboration by the OEM is one aspect that HP is touting for Itanium. "There are over 20 companies investing in this thing and that's a very different model than one architecture per company," says Barry Crume, HP's product line business manager for Itanium systems. "That means there is a lot of shared R&D [research and development] that is going into the basics of the architecture so there is a lot more money left to differentiate by vendor."

Crume cites SGI's decision to focus on developing Itanium systems for the Linux marketplace as an example of the type differentiation that will be seen in the future. SGI plans to launch the Silicon Graphics 750 system for Linux in July.

While the Itanium systems are hitting the streets in 2001 with early pilot programs, the next generation system, code-named McKinley, available next year, is expected to really take the new chip to new heights. "This [year's Itanium] is a widely deployed test vehicle," says Partridge, who believes people that use the new systems are simply preparing for the arrival of McKinley next year.

Crume also sees benefits down the road based on the early Itanium efforts. "Within a few years the 64-bit aspect of Itanium and its extra parallelism will work its way down to office desktops," he says. "We're not starting there because those aren't the people demanding the extra performance. This performance can be tough to get at it for engineering applications. We need to focus on it. It doesn't just come for free."–*patrick coleman, freelance writer*

SAN Vendors Try Cooperation

For all their benefits, storage area networks (SANs) are plagued with a reputation for being near impossible to work with. For some storage customers, the headache that comes with configuring different SAN components to work together ultimately outweighs the benefits of setting up a SAN in the first place.

This June, however, six prominent members of the Storage Network Industry Association (SNIA) made fledgling steps toward changing that by launching the Supported Solutions Forum, or SSF. As its goal, the SSF will aim to pave the way to greater interoperability between SAN components and cooperation between the vendors that make them.

What analysts really appreciate about the SSF is the cooperative support agreement that members must abide by.

Founding members of the SSF include Fibre Channel network equipment vendors McData Corp., Broomfield, CO; Brocade Communications Systems Inc., San Jose, CA; storage vendors EMC Corp., Hopkinton, MA; IBM Corp., Armonk, NY; Compaq Computer Corp., Houston, TX, and Hitachi Data Systems Corp., Santa Clara, CA. The SSF has indicated that its membership is open to any other SNIA vendors as well.

The SSF's first order of business was to announce two specific "Supported Solutions," or qualified SAN configurations made up of one another's products: Data Zones for either McData or Brocade switches that allow specific configurations of SSF storage vendors' products to coexist on the same Fibre Channel network. Zoning, explains Don Swatik, vice president of information sciences at EMC, assigns specific ports on the switch to ports on the storage subsystem, independent from other ports assigned to other storage resources.

Before establishing data zones, hooking up competitive storage to the same switch "might have created problems," says Swatik. With Data Zones, "we basically each have our own backyard to play in."

Going forward, the SSF will attempt to evolve SAN storage systems so they don't just coexist on the same SAN, but interoperate and share data across zones, says Swatik.

Nice as Data Zones may be, what analysts really appreciate about the SSF is the cooperative support agreement that members must abide by. "We think that 99% of the stuff that comes out of consortiums is crap," says Steve Duplessie, senior analyst with The Enterprise Storage Group, in Milford, MA. "But this one is beautiful, because it makes things easier for the end user."

Under the SSF's cooperative support mechanism, if an end-user calls a vendor, but the vendor finds that the problem is not theirs, that vendor is prohibited from passing the customer off–"they never leave the loop," says Duplessie. "They can't pass the buck anymore, no matter whose problem it is."

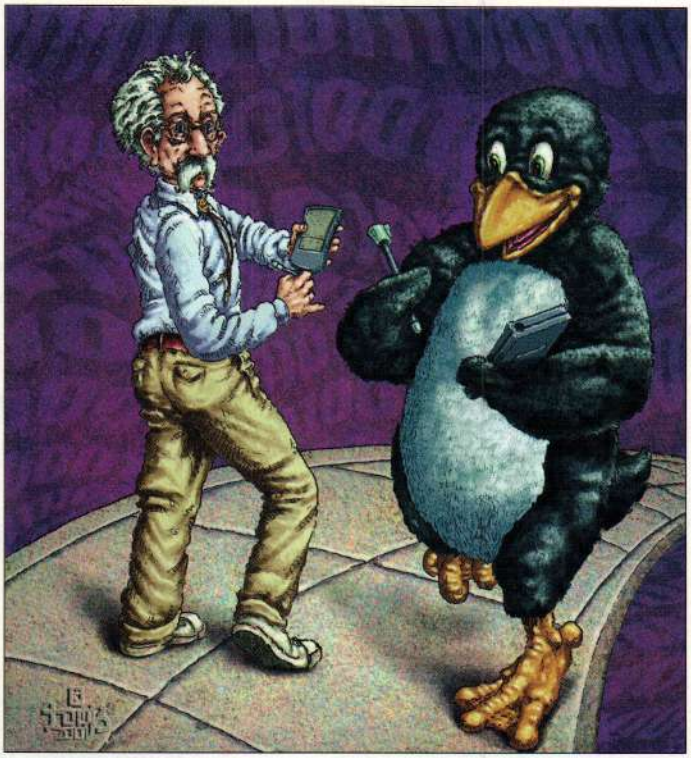
Other analysts agree this is an important step. "One of the problems plaguing SNIA is that it can push forward interoperability from a technical standpoint, but when members are out there publicly battling and blaming each other it tends to cloud the issue. This really brightens up the sky for storage networking," says John Webster, senior analyst, Illuminata, Nashua, NH.

Vendors are equally enthusiastic about SSF's cooperative support mechanism. "This is a huge thing," says Swatik. "In the past, we had to do interoperability testing on our own–without the other vendors' help. At least now we have a number to call."

But if SSF is to be truly effective, more vendors need to climb on board. "There were some very conspicuous absences, namely Sun and HP" says Duplessie. "If I were a customer, I would demand that they be a part of it, because that way they're not oddman out if something goes wrong."–*ab*

Ask Mr. Protocol

by Michael O'Brien



TOM BARRETT

Inside this package you'll find one Agenda VR3, the world's first Pure-Linux PDA.
– Agenda VR3 Quick Start Guide

The utility performs the following: Replaces the current ROM image on your device with an updated version.
– Compaq Support Web page

Sharp has announced a new PDA for the non-Japan market which features an embedded Linux operating system and a Java application environment.
– News story at LinuxDevices Web site

Mr. Protocol and Linux to Go

Q: Good heavens. Is that Gregorian chant I hear? This is the last place I expected to hear that.

A: Hey, now. Although we are no more than ordinarily cultured, we're not complete numps around here. But as it happens, that's just the alarm on Mr. Protocol's PDA. It's that Palm XXIII over there, the one he's had for the last few months. All white, triple crown and crossed keys on the screen cover, holographic angel announces he hours? What gets me is the way the sound seems to come from all corners of the room. And that bass! Haven't heard a good 64-foot diapason like that outside of a cathedral. Well, felt it, really, as a vibration in the chest cavity. It certainly beats THX all hollow. It's amazing what a company can do when it gets an influx of money. Who knew the Vatican was into high tech? It's just disgusting, though, the advertising that surrounds Palm Sunday these days.

I never mess with Mr. Protocol's stuff, though. It just makes me very nervous. Instead, I have been looking at PDAs that are a little more readily available and don't have ineradicable weekly to-do reminders about confession and such.

I might claim, as Will Rogers did, that all I know is what I read in the paper, such as the Palm VIIx now being available for $99, including a rebate, if you are willing to shell out about $24 a month, minimum, for wireless service. I'd be lying, though, since I also have Mr. P. available for an information channel straight into computational irreality– videlicet, his Palm PDA, among other things. But what use is a PDA that's too scary to pick up? Instead, I have been looking at some others. A Palm, after all, is rather a minimalist PDA, in terms of computational power and OS support. The two things that win big in the Palm arena are the sheer bulk of third-party software, and the form factor.

As Mr. Protocol's amanuensis, I take an extremely conservative view of computer equipment. I have to. If I get anything too advanced, Mr. P. just grabs whatever it is right after breakfast, and by dinnertime, it's unrecognizable. Remember how I said his computer looks something like a druid altar with a Bose-Einstein condensate in it? Well, the Celtic kick continues. He wandered by my regular old MH mail handler the other day when I wasn't looking, and when I tried to read my mail, I didn't have a text window any more. Instead I had a 3D graphic of a druid holding a forked stick, who shrugged his shoulders and said, "Diabhal scéal!"

It took me a while to get that translated as, loosely, "Not a thing, boss!"

Entrancing as it might be to read my mail via spoken commands in Irish, I made him put it back the way it was. He was hurt. He couldn't understand my lack of class, and ignorance of my roots. I threatened to plant him in some roots, and it went downhill from there. I had to reinstall MH myself, always a pleasant prospect, and delete about 80 GBs of God-knows-what that he'd put in there.

Feel the Power!

UltraSPARC™ III

Hang on to your seat and feel the power of the 2x or more increase in performance. UltraSPARC III technology brings high-speed graphics, blistering data transfers and faster everything to your system.

100% SUN MICROPROCESSORS

As a Sun® Master Value Added Integrator, Tatung uses 100% Sun UltraSPARC CPUs and motherboards. Rely on Tatung's 12 years of design and manufacturing experience with SPARC-based systems and have them build a system for you.

Visit our website, www.tsti.com,
or call 1-800-659-5902
for the office nearest you.

ULTRASPARC™ DRIVEN

Point to tsti.com

TATUNG®
SCIENCE & TECHNOLOGY, INC.

Sun, is a registered trademark of Sun Microsystems, Inc. in the United States and other countries. Tatung is a registered trademark of Tatung Company. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the United States and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc. All other products or service names mentioned herein are trademarks of their respective owners.

As it is, the monitor still has this green tint I can't get rid of.

All of this has led me to a much more conservative view of my own computational environment.

It's also led me to things that can be easily hidden.

This means that I look at a lot of PDAs. In particular, my main portable computational device remains the Apple Newton 2100. It's got an actual screen and is reasonably speedy. Moreover, if the batteries go flat, it doesn't take a complete dump the way a Palm does. It does suffer from the attribute that killed it, though: it won't fit in a shirt pocket unless you wear shirts tailored for Paul Bunyan. It does, however, have a wickedly advanced OS and application environment. Newton applications are written in an advanced language called NewtonScript, an object-oriented (OO) language derived from Self, which in turn was derived from Smalltalk. It is almost impossible for a NewtonScript program to go amuck in such a way as to freeze the machine, unless it runs the Newton out of resources. Palm applications are written in C, so when a Palm application goes amuck, it generally either freezes the machine or causes a system crash and reboot.

So, what's around in the way of pocket-sized PDAs? Mr. Protocol is glad you asked. Just please don't listen to a thing he says, or soon you too will know the joys of being awakened by the Angelus at six in the morning.

PDAs are taking off. There are a lot of them on the market in all price ranges. All have learned from Apple Computer's expensive mistake with the Newton, and almost all now fit in a shirt pocket. Beyond that, there's not a lot in common between them.

Some have keyboards, though in some cases, like the Blackberry wireless e-mail/pager devices, you have to type with two thumbs on a keyboard with keys the size of rice grains. All Psion devices have keyboards, and their own operating system, EPOC-32, has been adopted by the Symbian Alliance for use in wireless devices. Mr. Protocol hasn't actually heard anything bad about the Symbian Alliance, but you'd think we'd be seeing some EPOC-based wireless devices on the market by now. Ah well, they say Bluetooth isn't dead either, just slow to market. We'll see.

Stylus and Style

Most PDAs, though, use styluses, with either a tiny onscreen keyboard, or some sort of handwriting recognition. Which of these is better is sort of a push. Many people are proud of how fast they can write Graffiti characters on a Palm device, but then, this may just be cognitive dissonance. Others just prefer the more natural feel of writing than pecking things out on a tiny keyboard with a stylus. In either case, data entry via a stylus is much slower than typing on a keyboard, at least for most people, so the first thing that stylus-based PDAs have to do is to make a little data entry go a long way. If possible, people should be able to pick commonly entered items off a list, instead of laboriously entering the whole thing each time. Be it noted here that the Newton had this down cold: most of its applications remembered the last several things entered into each field, and allowed you to pick them from a pop-up menu.

Beyond that, it's more a matter of personal taste than anything else, at least for ease of use. However, there are some more important things that become apparent once one has worked with a PDA for some time. Let's take an in-depth look at a couple of entries in the market to see where they take us.

At the high end of the market, we have fancy little numbers such as the Compaq iPAQ 3635, the newest entry in the handheld PC field from Compaq. This device, which goes for about $600, comes preloaded with Windows CE 3.0. And it's at this point that the religious wars start. Let me state it here: I am a diehard Newton user, but believe it or not, Windows CE 3.0 is not as revolting as I had expected it to be. A number of major improvements have been made that increase its usability to the point where you do not actually want to throw it against the wall after five minutes. To be sure, it's still Windows, and that's bad. But it's not as bad as it could be.

Windows CE was designed at all levels to leverage off the general familiarity of the Windows desktop interface. This turned out to be a major blunder. Windows is an indifferent desktop interface, but it is an absolutely unusable interface on a handheld. Oh my and oh brother, people hated it, even people who were predisposed to praise anything Microsoft does. Mind you, people who do this quickly attain positions where not praising Microsoft will cost them their jobs, so when they saw what Windows CE was like, they smiled, said how nice it was that it was just like the Windows desktop, then snuck out at 5:01 p.m. and bought a Palm in some suburb where no one would recognize them.

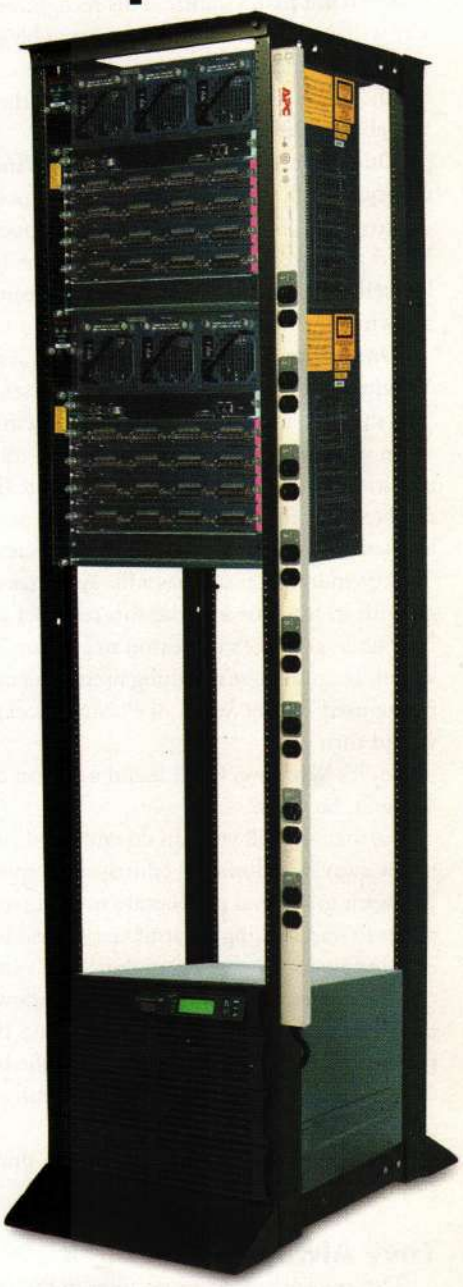
One thing the Apple Newton got right is that a handheld is not a desktop. Not only is a pen-based interface radically different, but the applications differ as well. For one thing, the number of things that you actually want to do on a handheld is much smaller than what you'd do on a general-purpose desktop. There are exceptions to everything, but mostly, people use PDAs the way they use a Franklin organizer: to keep contacts, to-do lists, notes and calendar items. Even when they connect it to the Internet, the small size of the device and the pen-based nature of the interface mean that the mail and Web browsers will look quite different than they do on a desktop.

To the extent, then, that Windows CE 3.0 does NOT look like Windows 98, it's an improvement. In fact, the Compaq iPAQ 3635 can do some things that till now only the Newton could do. Unfortunately, the Newton still does them better, but this may not be the fault of Windows.

The iPAQ 3635 brings back the ability to write cursively on the screen and have the device recognize one's handwriting and turn it into text. Unfortunately on the iPAQ this amounts to a party trick. Because it's supported only in the Notepad, you have to explicitly request that each batch of handwriting be recognized, and the recognizer only does a job about as

Some VoIP conversations should be interrupted, but never by power problems

APC provides all the components necessary for an end-to-end power protection solution for the VoIP environment.



The 7 pieces of the VoIP availability puzzle

- **Clean, continuous power** as well as **"ride-through" power** during brownouts, surges and spikes.
- **Extended back-up power** in the event of an extended power outage.
- Redundant, hot-swappable and scalable components to **allow growth** as well as **service without interruption.**
- **Instant notification of critical power/UPS issues.**
- Ability to remotely control selected power outlets in order to **reboot hung switches.**
- Ability to **ensure optimal temperature and humidity** within remote closets.
- Ability to keep track of and **maintain health of power protection systems** across the WAN, over time.

APC provides all the components necessary for end-to-end power protection solutions for the VoIP environment – visit ***apcc.com/buy/*** and see what Legendary Reliability™ can do for your business.

Symmetra® RM

The new Symmetra RM puts the high availability of the proven and patented Symmetra® Power Array™ technology in a rack-mountable form.

Through the included Web/SNMP Management Card, you can monitor and configure your APC Symmetra RM to shut down and reboot your systems, receive e-mail alerts and view the event log.

Remote Monitoring

APC monitors all UPS parameters, tailored to your desired response. Regular UPS parameter and event reports are issued with event frequency, duration, and resolution, offering immediate enhancements to your investment.

MasterSwitch™ VM

Provides the ability to monitor the current draw and set alarm thresholds, based on customer requirements, while still providing the remote on/off/reboot capabilities found in the MasterSwitch series. In addition, it mounts vertically, requiring zero U of valuable rack space.

Environmental Monitoring Card

Works with your APC Smart-UPS® or Matrix-UPS® to monitor ambient temperature, humidity and other environmental conditions.

PowerChute® Inventory Manager

An invaluable software tool for anyone with a large number of APC UPSs spread across a wide geographic area. Via SNMP-enabled APC UPSs, schedule the software to gather information from the UPSs, then select any one of the eleven predefined reports.

Other APC products for the VoIP/Rack environment:

- **KVM Switches** provide one centralized control point for up to 64 servers.
- **ProtectNet** rack-mounted data-line protection.
- **PowerNet Manager** collects UPS/power status information for fast problem diagnosis
- **Cable Interface Kits** provide direct communication between UPSs and desktops, workstations and servers.
- **2-Post Racks / 4-Post Open Frame Racks**

Tivoli · Solaris Compatible · Designed for Microsoft Windows 2000 Professional, Windows NT Workstation 4.0 · HP OpenView Premier Partner

Symmetra RM is easily manageable with the industry-leading network platforms.

*By utilizing **APC's PowerNet for CiscoWorks2000**, which integrates APC's power management software with CiscoWorks2000, Cisco customers now can easily manage APC power protection and network power control devices from the same Web browser as Cisco equipment.*

Cisco Systems Verified

INFORMATIONWEEK 500

APC was named to the 2000 InformationWeek 500 ranking of the top IT innovators (09/11/00).

APC®
Legendary Reliability™

Enter to win NEW Server room air conditioning unit from APC!

All entrants will receive a FREE Enterprise Availability Kit.

To order: Visit *http://promo.apc.com* Key Code a838y • Call 888-289-APCC x2041 • Fax 401-788-2797

©2001 American Power Conversion. All Trademarks are the property of their owners. APC1H1EF-US • PowerFax: (800) 347-FAXX • E-mail: apcinfo@apcc.com • 132 Fairgrounds Road, West Kingston, RI 02892 USA
Microsoft, Windows, are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

good as the Newtons did before handwriting recognition got fixed in Newton OS 2.0; that is, it's not very good at all.

On the other hand, for character-at-a-time recognition that works in any application, they do have a recognizer that's a lot easier to use than Graffiti. This recognizer deals with individually written letters, but the letters can be written normally, i.e. not in the funny Graffiti alphabet Palm users learn to love (or hate). This is a vast improvement, and the recognizer works just about 100% of the time.

One of the biggest advantages of Windows CE is the ability to exchange Microsoft Office documents between the desktop and the handheld. There are pocket versions of Word and Excel provided, but as others have noted, the really useful one would be Pocket PowerPoint, and that ain't anywhere in sight.

One of the big disadvantages is the survival of some Windows utilities, such as the File Browser. Where the iPAQ puts things, and what's in those files, doesn't seem to have a lot to do with logic. The Notepad and the rest of the PDA-style applications keep all their data in private little databases. In the Newton, these databases weren't so private, and the PDA applications could interoperate. There was no file system per se. Well, Windows CE still has a file system, and a File Browser to go with it, too, but so far as this reviewer can tell, it's spectacularly useless. There's no reason to go grotting around in the file system because there's nothing useful you can do there. Perhaps if one used Pocket Word or Pocket Excel a lot, something would turn up.

So, it's Windows CE 3.0 and a person could almost stand to use it. So what?

So, that's not all you can do with an iPAQ, is what. You can throw away Windows CE entirely and replace it with Linux if you want to. Several projects are working on this. A couple of places to start looking are at `http://familiar.handhelds.org` and `http://www.linuxdevices.com`.

These ports run a very small X Window server, drawn from the XFree86 server familiar to Linux and BSD types. Well, X runs quite well on desktops, but who the heck wants all that window manager top border schlog taking up precious room on a handheld?

Good question and just a little too practical for Mr. Protocol to care that you asked.

They Always Come Back

This movement seems peculiar in the extreme. The immediate reaction of most people with any sense is, "This is crazy." They're going to spend man-years on this, and come up with something that's way more expensive than a Palm and not quite as good.

On the face of it, this is hard to argue with. But there are several points in favor of this work, the first couple of which should be familiar to people who've been around a while.

1) It's neat. Hard to argue with this. I mean, you could, but there's no point.

2) It's all open source. You can hack your PDA! That's never been possible before. (But why would you want to? See argument for #1, above.)

3) The actual operating system running on a PDA doesn't really matter, so long as it fits on the hardware, and Linux has been around a lot longer and is a lot more stable than anything currently running on handhelds. What matters is the interface that's presented, and X is flexible enough to let us "Do The Right Thing" when we get around to polishing that part.

Well, this is a funny one. It's true and it's not true. One reason Linux is stable on the desktop is that it has real hardware memory management backing it up, so if an application goes haywire, it can't crash the whole system. As we've seen,

Even on the desktop, Windows has only recently come to grips with the notion that real hardware memory management is a Good Thing.

Palms go belly-up regularly. Newtons would too, except that Newton applications are written in a language where you have to work really hard to get hold of memory you're not supposed to have. With Linux on a handheld, it all depends on the hardware, and if any current PDAs have hardware memory management, I haven't heard about it. They might, but let's face it: even on the desktop, Windows has only recently come to grips with the notion that real hardware memory management is a Good Thing.

On the other hand, if you are not content to just live with whatever Microsoft or Palm give you in terms of software, this is about the only game in town unless you want to go totally retro and buy a Newton off of eBay. (It should be noted here that the Newton mailing list is seeing an increasing number of people who are doing exactly this: having given up the dead Newton platform for a Palm, they use the Palm for a year or two, and after losing everything one too many times when the batteries go dead, or after experiencing the sheer frustration for one more time than they can stand, they go and buy a Newton and scream, "Why did I ever leave?" Go figure.)

The real secret to creating a PDA which is brilliantly easy to use, as opposed to learnable after much fighting and effort, is to support the interface seen by the user all the way down to the hardware. Few PDAs do this, mostly due to the laggard state of software engineering. The Palm line, for example, is programmed in C, which means that errors in an application usually cause the Palm to crash. A C program on a Newton will do the same thing, but the NewtonScript application programming language prevents applications from being able to do this. A mature OS, such as Linux, could allow such support, but doesn't necessarily provide it.

Now, the only reason you can hijack an iPAQ into running

Linux is the operating system in an iPAQ is not contained in ROM. It's contained in flash memory, and flash memory can be rewritten. This is a benefit. The Newton used flash memory to hold absolutely all data, so that even if you took the batteries out of a Newton and let it sit for an hour or a year, all your data would still be there. The iPAQ under Windows CE isn't quite this nice. It also dumps your data if you let the battery run all the way down, but on the other hand, the operating system, being in flash memory, will still be there. And since the OS is in flash memory, it can be overwritten when it's time to upgrade the software. This is actually pretty neat. Download an upgrade package from the Compaq support site, put the iPAQ in the synchronization cradle, and let 'er rip. Pretty soon you've got a new OS.

Of course, what you write into flash memory needn't be Windows CE. It could be Linux, and that's just what these people are doing.

And that's not all. Some people actually believe you can make a buck doing this sort of thing. There is a company that is selling PDAs on the cachet of Linux alone.

Agenda Computing has, as this is being written, just started to deliver the Agenda VR3, a very inexpensive PDA at $250. It runs Linux and X Window out of the box. You can synchronize it to either a Linux desktop or a Windows desktop. Is it a usable PDA? Surprisingly, yes it is. And they got some things right as well. Like the Newton, it uses flash memory for data, so you will not lose everything when the batteries die and go flat. It keeps the OS in flash memory, too, so you can upgrade it. And a good thing, too, because let's face it, this box is just a bit raw, like any radically new device. But it does work, and is a useful PDA to boot. And although it runs X Window, you wouldn't know it till you peeked under the hood: they have got a window manager that gets rid of the funky bars at the tops of X Window, and their applications run full-screen. There are bugs here and there, to be sure, but nothing that further development and reflashing won't cure. And I did something quite unusual around this place. I pointed the screen at Mr. Protocol and hit the reset button and watched the look of amazement on his face as the device showed a Linux boot sequence in the palm of my hand.

Is this a genuinely good idea? Probably not, but in all fairness, this is grounds for a ferocious debate. Will Agenda Computing survive? Probably not, but then, most new companies don't. Is the Linux PDA movement worth watching? Oh, my, yes. ⊷

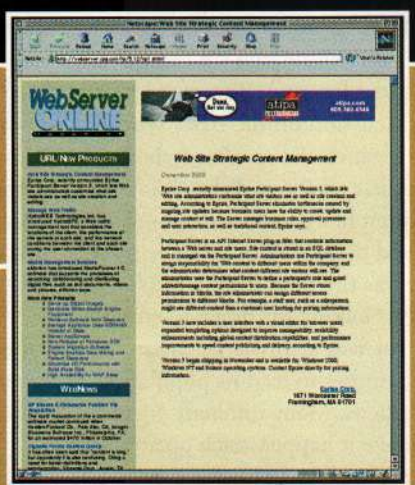
***Mike O'Brien** has been noodling around the UNIX world for far too long a time. He knows he started out with UNIX Research Version 5 (not System V, he hastens to point out), but forgets the year. He thinks it was around 1975 or so.*

He founded and ran the first nationwide UNIX Users Group Software Distribution Center. He worked at Rand during the glory days of the Rand editor and the MH mail system, helped build CSNET (first at Rand and later at BBN Labs Inc.) and is now working at an aerospace research corporation.

***Mr. Protocol** refuses to divulge his qualifications and may, in fact, have none whatsoever. His email address is* amp@cpg.com.

ATTENTION

WEB SITE MANAGERS...



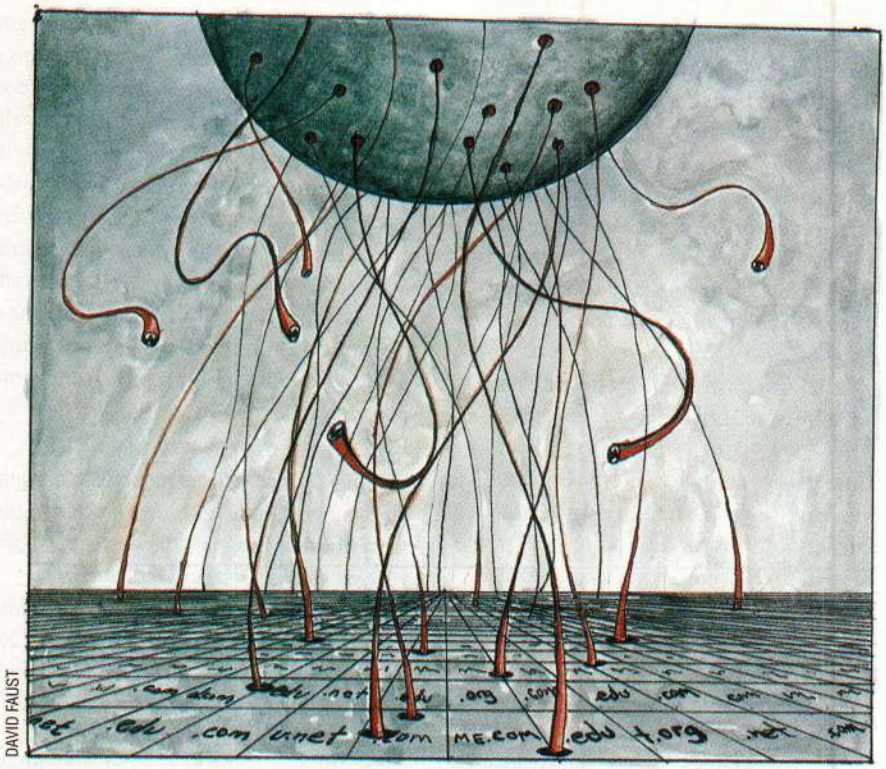
WebServer Online Magazine provides technical personnel with the essential news and new product information they need to develop and manage corporate Web sites.

Come to webserver.cpg.com to see *WebServer OnLine's* monthly news and new product information.

WebServer ONLINE MAGAZINE

http://webserver.cpg.com

by Peter Collinson, Hillside Systems



Domain Names

Many people find it odd that there is no one "in charge" of the Internet. Recently, I heard a radio discussion on the BBC World Service in which the speaker (whose name I didn't catch) said that human society often has complex structures unregulated by officialdom. His primary example was the food supply to New York City. Nobody sits down to ensure that food arrives in any large city in sufficient quantities to feed its population, it just happens. Many different organizations make it happen, each operating for its own benefit. I am sure that it's possible to come up with other examples. The Internet was bootstrapped by the U.S. Government but now has a life of its own, run by many separate organizations whose self interest keeps the whole thing operating.

One area of huge expansion in recent times is the registration and use of domain names. Everybody with a Web site is encouraged to have at least one domain name. Most of the sites that sell them actively encourage you to buy more than one. Since humans prefer to deal with names rather than numbers, the basic function of the name is to act as a human readable tag that addresses a specific machine. The mapping of names is managed by the DNS (the Domain Name System or Service) a distributed database that underpins the Internet.

The familiar dot-separated naming mechanism provides us with a way of constructing unique names for a particular machine and is used to provide us with a way of walking through the name space to find the address of that machine. For example, let's look at the name `frodo.cpg.com`. It is navigated from the right-hand end. We first find the name server that knows about the "top-level" domain `com`, then in that database find the server for the `cpg` domain, then finally we look for `frodo`.

Initially, various "top-level" domains were established, `.com`, `.org`, `.edu`, `.net`, `.mil` and `.gov`. This initial set is currently being extended. Countries outside the U.S. tended to use their standard two-letter country code as their top-level domain.

A key idea of the DNS is that of *delegation*. Once a domain is established, such as `.uk`, then the names within that domain can be run by an organization politically, geographically or economically closer to the users of the domain. When the `.uk` domain was established, the powers-that-be, who were academics, maintained and ran the machines that supported the `.uk` name space. They split the `.uk` name space into `.ac.uk` for universities, `.co.uk` for companies and `.gov.uk` for government.

Some years ago, the whole U.K. operation was dumped in the lap of a not-for-profit company, Nominet Ltd. Some parts of the name space were delegated to other organizations, for example, `.ac.uk` is now run by part of the academic infrastructure. Nominet runs some parts itself. My domain `hillside.co.uk`, for example, is registered with Nominet. In turn, they delegate the

running of my domain to me, so I can create names in it as I wish by altering files on my machines.

The story in the U.S. is similar, as name management moved from being run by volunteers and funded by government research budgets into the hands of a single commercial registrar. It moved on from there in late 1998. The general management of names on the Internet as a whole is now run by a not-for-profit corporation called ICANN (Internet Corporation for Assigned Names and Numbers). The best way to explain this organization is to steal a paragraph from their Web site:

"Formed in October 1998, the Internet Corporation for Assigned Names and Numbers (ICANN) is a non-profit, private-sector corporation formed by a broad coalition of the Internet's business, technical, academic, and user communities. ICANN has been recognized by the U.S. and other governments as the global consensus entity to co-ordinate the technical management of the Internet's domain name system, the allocation of IP address space, the assignment of protocol parameters, and the management of the root server system."

The "root server system" is the set of machines that manage the top-level domains. ICANN sees itself as a technical coordinating body; its job is not to "run the Internet." ICANN has also been responsible for expanding the number of registrars from one–Network Solutions–to many, each able to create domains in the top-level name spaces. Some new top-level domains: `.biz` and `.info` are being created to expand the global name space.

Confusion

Although the people who manage some domains have established rules for who can register in that domain, names have mostly been allocated on a first-come, first-serve basis. While this seems an equitable policy, it's certainly fostered confusion and obfuscation. It's also lead to quite a bit of annoyance in the business world, where names are money.

A good example of confusion is illustrated by the domain `whitehouse.gov`. This, as you might guess, is the official Internet address of the building in Washington, D.C.–the Executive Office of the President of the United States. The Web site found at this address contains interesting facts about the latest incumbents: Spot, Barney, India the cat, and a few humans whose names you might know.

Well, when looking for the White House, you might think to type the address `whitehouse.com`, perhaps by accident, or in complete innocence. If so, you will find a brightly lit site containing a different set of incumbents, having names like Mandi, Celeste and Sandy. If the truth be told, you don't get to learn much about them from the site. You are simply presented with detailed photos of their particular physical attributes and some statistical information about them. You will have no doubt that this is certainly not the official site of government.

It's perhaps more confusing to type `whitehouse.net`, which is certainly another plausible address for that same building. Now, this looks very similar to the `.gov` site and is executed technically with more skill. It seems to have several different front pages (use Reload on your browser). A closer inspection shows this `.net` address hosts a spoof of the "real" `.gov` site.

At the time I am writing, the `whitehouse.net` site contains a page of reader's responses that displays a letter from someone complaining about the `.com` site as if Spot, Barney and India could do something about it. The letter illustrates the confusion. The author is concerned that the soft porn site is only three characters away. She is worried about allowing her kids to see the images by accident, and this is a common concern. She can, if she wishes, subscribe to a Net monitoring service that will filter out this site for her.

Of course, she wants to ban the site. Our joint puritan heritage is strong on making a lot of noise about banning things. The U.S. is spared some of the furor due to the strong sense on the freedom of speech that is embedded in the psyche of the population, which in turn derives from U.S. constitutional rights. Here, in the U.K., the "I don't like it and don't want you to have it" lobby is often loud on a wide range of topics. After all, the Internet is currently a tag word that sells papers. Incidentally, unclothed well-endowed females are deemed to sell papers too. It's certainly the case that U.K. children can see images of unclothed female torsos in at least two national tabloid newspapers, should they be interested.

Regulation

Please don't get me wrong. I am not arguing for regulation of anything. I don't think someone should be able to sit in judgment and say that `whitehouse.com` or `whitehouse.net` should not exist. Politicians are all trainee comedians, and poking the fickle finger of fun at them is pretty much all they seem to be good for. They are so boringly bland, and some laughter at their expense may help us all feel somewhat less scared about what they can do to us, should they be mindful.

However, the confusion engendered by the free-for-all with domain name registration can be considerably more meaningful for companies, large and small. Companies all spend money on promoting their name or trademarks, and we have national and international systems in place predating the Internet by some years that are supposed to protect that investment. Name protection is generally a good thing as it mostly allows us to be certain when we buy a branded product, we are getting what we have paid for.

Many companies have taken to registering all possible permutations and combinations of their names in as many top-level domains they think fit. The intention is to prevent clone sites from being created using the names. There are people out there who will register names to annoy certain companies, people who register names that are close to a typo or a common misspelling of other names (see `http://www.macdonalds.com`), and people who will register names as an investment. They hope that the company will come along, want the name and pay through the nose for the privilege.

Some things are being done about all these problems. The new `.biz` domain, for example, is starting life allowing for trademark registration. In the U.K., Nominet only allows the registered company name to be used in names in the `.ltd.uk` and `.plc.uk` domains. Since we have a national registration of company names, this guarantees uniqueness of names and ensures that they will be used appropriately.

What is What?

How can you determine whether a site you are looking at is run by the organization that purports to own it? Here's a related question: How can I find a contact for a particular domain?

Most of the level domain administrators run a `whois` server to provide information about registrants of names in their domain. The problem is that these servers only contain information about the domain they support and there is no way of asking one server to find another. So your first job is to find the `whois` server for the top-level domain you are seeking.

If the domain you are seeking is in `.com`, `.net` or `.org`, then the `whois` server is `whois.internic.net`. If it's another domain then you may have trouble guessing. There are now several Web sites that will search the appropriate `whois` server for a domain name. On Webmagic's site `http://www.webmagic.com/whois/index.html`, you'll find a form where you can type the domain name, select the top-level domain and the system will interrogate the correct server for a response. It will tell you whether the name is in use and give you the option of seeing the information the server contains. The site `http://www.alldomains.com` does a similar job. Run by the same company, `http://www.allwhois.com` searches the server and shows you its output directly.

I have to confess to being nervous about using Web-based search systems when looking for a new name to be registered for a client. It's entirely possible for the domain names that you type into these forms to be logged, and a pre-emptive registration can then be done, perhaps before you can get back to your client and negotiate about whether they want the name or not. I am not accusing the sites mentioned above of doing this, I am merely stating the possibility.

I prefer to use the `whois` client that's found on my UNIX system to interrogate a server. If you type:

```
$ whois cpg.com
```

you'll get output shown in Figure 1.

The `whois` command interrogates a default `whois` server, and you may have an old and defunct server compiled into the program. You can give the program an explicit server to talk to by using the `-h` option. Your `whois` client will need to connect to `whois.internic.net` to obtain the information above.

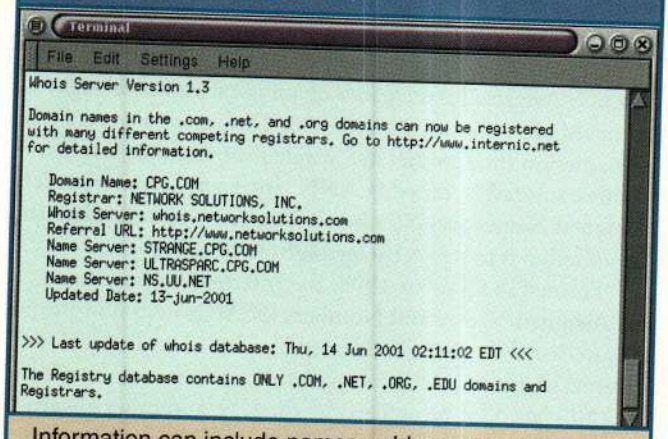
In fact, for the U.S. domains in the default server, you have to give a second command to find the full information about that domain. Take the server name from the `Whois Server:` line in the initial output shown in Figure 1 and type:

```
$ whois -h whois.networksolutions.com cpg.com
```

and you get the full output for the domain. You can obtain several names and addresses: the owner of the domain and the addresses of technical, administrative and billing contacts. The record will also give the names and IP addresses of the name servers that support the domain.

While I was doing research for this article, `whois.internic.net` disappeared from the Internet for a bit. I logged into a friend's FreeBSD machine (that happens to be in California) to see whether it was my ISP that was at fault. It wasn't. However, I found that the version of `whois` on the FreeBSD system has been updated to cope somewhat better with the current situation. It will automatically perform the second `whois` lookup for you.

Figure 1. *Whois* server


Terminal
File Edit Settings Help
Whois Server Version 1.3
Domain names in the .com, .net, and .org domains can now be registered
with many different competing registrars. Go to http://www.internic.net
for detailed information.
Domain Name: CPG.COM
Registrar: NETWORK SOLUTIONS, INC.
Whois Server: whois.networksolutions.com
Referral URL: http://www.networksolutions.com
Name Server: STRANGE.CPG.COM
Name Server: ULTRASPARC.CPG.COM
Name Server: NS.UU.NET
Updated Date: 13-jun-2001
>>> Last update of whois database: Thu, 14 Jun 2001 02:11:02 EDT <<<
The Registry database contains ONLY .COM, .NET, .ORG, .EDU domains and
Registrars.

Information can include names, addresses and domain owner.

Also, and this is a huge step forward, the FreeBSD `whois` automatically finds the appropriate `whois` server for the domain you are seeking. A special `whois` server run by Centergate Research Group, LLC, makes this work. To look up a U.K. domain name, make a query like:

```
$ whois -h uk.whois-servers.net hillside.co.uk
```

To create a query using `whois-servers.net` take the top-level domain and prepend it to the server name. If you type this command, you will be given the information for my domain that's presented to the world from the main UK `whois` server at `whois.nic.uk`. You'll find that the information is different from that given to you by U.S. servers. You'll only get a registrant name, the company that registered the name in the form of a tag, and the name server information. It's quite typical that different `whois` servers will present different information.

It's simple to create a script that will do the work for you and use the older `whois` client that should be available on your machine:

```
#!/bin/sh
cmd=/usr/bin/whois
if [ $# = 1 ]
then
 # get the last component of the name
 tld=`expr $1 : '^.*\.\([a-zA-Z]*\)$'`
 # run the whois command
 $cmd -h $tld.whois-servers.net $1
else
 # call the real whois if we have more than
 # one argument
 $cmd "$@"
fi
```

I want to put this in my private bin directory and call it whois, so I arrange that if the script is called with more than one argument, I will call the standard command with those arguments. If I have one argument, then I am assuming that it's a domain name. This is not quite complete because whois servers can contain information other than just domain names, and I am neglecting the possibility of this.

If I have one argument, say cpg.com, then I want to pull off the last component of the name, com, to prepend to the server name. I've chosen to do this with the pattern matching capabilities of the expr command. There are many other ways. The expr command is run inside back quotes so that its output is captured and placed in the shell variable tld. There are three arguments to expr. The $1 replaced by the first argument to the shell script, the colon says "do a pattern match that matches the regular expression in the third argument against the first argument." There's some magic here that lets us pull out part of the source string and output it. Let's pull that regular expression apart (see Table 1).

So if I am matching against cpg.com, then the match string will pick out com and expr will output that string. It will be loaded into the tld variable and injected into the command on the next line. This regular expression could be made more robust and it is possible to test the value of tld to see if anything has been matched.

How Useful is *whois*?

It all depends on the amount of information you can retrieve. The U.K. server hardly gives you any information about the name registrant. On talking to Nominet, I find that they have to operate within the provisions of the Data Protection Act in the EU and are constrained by that law. However, they do offer a free dispute resolution service and will help to make contact between people should that be needed.

In the U.S., the whois servers supply more information. This can be a good thing, and it can help you obtain a contact for whoever owns the name. It can also be a bad thing. There's a huge temptation for people to use the e-mail information the servers contain to send bulk e-mail or even one-off advertising. Many whois servers will transmit a disclaimer that is often ignored. It's hard to prove the exact source used by the junk e-mailer to get their information.

You can certainly use the whois information to find some of the truth about the three White House sites. I say *some* of the truth, because it seems that none of the address information presented by whois is validated by the name registrars. They have historically depended on people supplying correct information. First, the information can be stale. It could have been hanging around for a number of years. Second, it's possible to register with a legitimate address, pay the bill, which acts as some form of validation, and then change the stored information to obfuscate your origin.

I had an incident recently where I was trying to find out who had registered a name that was essentially the same as the company name of one of my clients. The whois information for the domain was entirely bogus, and I used the DNS to find a contact in the reseller company who hosted the domain for the DNS. The person there did reply to my e-mail query, but said I needed a court order to obtain the contact information I was looking for.

After some advice from the people who had actually registered the name, and were directly responsible to ICANN, I complained to their compliance department. The name owner was forced to supply an address, but I don't believe that it's real. It's uncheckable since it's in Kuwait. The phone numbers are certainly bogus. It's all a dead end, which is unsatisfactory.

The real problem with the privatization of the domain registration business is that the actual registration has been placed in the hands of anyone with a Web browser and there are no checks on the data that is supplied. ICANN may regulate its clients, but actual name registration is often being done by organizations that are three (or sometimes four) steps away from ICANN. Pricing is so low there is no money to provide for support staff to validate the information. It's trivial to install cut-offs and blind alleys in the whois information so that it's not possible to find the owner of a name. ISPs and name registration companies hide behind the law and refuse to assist you, even if you have a legitimate query. It's a free-for-all where the unscrupulous flourish. I find it all rather worrying.

Table 1. Magic in the Script

^	matches the start of the source string
.*	followed by anything–this is being discarded
\.	followed by a dot; it has to be quoted with a prepended backslash to match a real "dot" in the source expression
\(	the start of match section; if the expression inside the match string is found, then the contents of this match section are output by expr
[a-zA-Z]*	The match section in this case will be matched by any combination of letters in upper and lower case repeated as many times as we need
\)	the end of the match section
$	the end of the string

Finally

As I've said, the DNS can offer a way of tracing the real owners of names. At least you can try to track the people who are hosting the names and it can help–sometimes. There are also two commands of interest–nslookup and dig. I obtained some of the information about whois from *DNS and BIND* by Paul Albitz and Cricket Liu, published by O'Reilly and Associates Inc., fourth edition (ISBN 0-596-00158-4). You'll find more information about how to interrogate the DNS in this book. ⊸

***Peter Collinson** runs his own UNIX consultancy, dedicated to earning enough money to allow him to pursue his own interests: doing whatever, whenever, wherever ... He writes, teaches, consults and programs using Solaris running on an UltraSPARC/10. Email:* pc@cpg.com.

by Henry Newman

COTEY COMES

Secondary Storage

This month's topic covers a broad set of secondary storage technologies in which change has been slower than in hard disk storage. Where disk densities have increased 2,406 times over the last 25 years (80 MB to 188 GB) linear tape densities have only increased 667 times (150 MB to 100 GB). Even including helical tapes and the densities, ratios do not match for hard disk size increases, which is 1,333 times. Today RAID is often used instead of a single hard disk so the effective storage on a channel is even higher. Optical densities are increasing at even slower rates, even though some significant breakthroughs have recently been made with blue lasers. The real issue is that transfer rates for tapes and optical have in no way kept up with these increased densities (see Table 1, Page 26).

At best the transfer rate has increased 19.2 times while RAID peak performance increased 133 times during the same period (200 MB/s full duplex). So we are almost an order of magnitude difference. These ratios are not going to change much anytime soon.

My good friends at Imation Storage Consulting have been providing me with a low-level media education over the last two years. This month we are going to cover both tape and optical media issues including: error recovery and rates, shelf-life, and media. Since I'm not a world-renowned expert like my friends at Imation, I asked them to review this month's article to make sure I got it right.

Compression

Data compression on tape is at best a tricky business. Vendors make claims that often do not match reality. As the old saying goes "your mileage may vary." A few things are universal truths. You will never get better compression than GNU *gzip* because no vendor uses a hardware data dictionary and hardware compression buffer as large as is used by *gzip*. If you have image data, compression is not likely to be very high because images are already compressed data format. If you are saving e-mails you can get 20-to-1 or greater with *gzip*. Optical does not support hardware compression, another truth. Binary data compression varies greatly and generally you must test a large number of files to get a statistically significant sample of your data. Last but not least, from what I have seen high-end tape drives have *significantly* better compression than lower-end tape drives. The amount varies depending on the data so I cannot give (no one can) you a number because it is so data dependent.

If you have better compression with high-end tapes even though the drives cost significantly more, the reduction in the amount of media might save money–at least in media costs. If using a library you might be able to save on slots and purchase a small library depending on the tape drives you chose and the compression for those drives. Don't count out high-end tape drives based on price alone. Be sure to consider your overall

media costs based on compression and the cost of a robot and number of slots needed.

I looked far and wide for a tool to measure tape usage for all common tape types. I was unable to find a tool, so if any reader has heard of such a tool please let me know and I will pass it along. To determine tape usage, you need to probe the tape drive and get information regarding the usage as the drive tracks it. You need to send a direct SCSI command to the drive and get the specific sense bits that the drive manufacturer uses to track tape usage. Every vendor uses a different set of bits and format. Some software products such as backup and HSM provide such tools, but no general tool for multiple drives and operating systems is available that I could find.

Media Types

To fully understand tapes and tape issues we need to fully understand the two different types of media–linear and helical scan–and the implications of those two media types for speed, density, reliability and cost. Unfortunately, you cannot have the best of all four.

With linear tape recording technology, data is written in tracks that run the length of the tape media. A large number of vendors make linear tapes: IBM 3590B/E, STK 9840/9940, Quantum DLT 7000/8000 and Super DLT, and many vendors sell LTO Ultrium and QIC. As you can see in Figure 1, the recording tracking for linear tape drives is parallel and determining the track alignment can easily be done with laser guidance.

In contrast, helical scan systems record data in diagonal stripes across the tape (Figure 2). This diagonal pattern allows data tracks to be overlapped for high data density.

A number of vendors make helical tapes: Sony makes AIT-1 and AIT-2 as well as DTF-1 and DTF-2. Other helical types include 8MM, Mammoth and Mammoth-2, 4MM (DAT), and STK Redwoods. For helical tapes the heads must be aligned to a greater degree of accuracy (Figure 2).

There are currently three types of optical storage: Magneto-optical (MO), DVD/CD-R, and Write Once Read Many (WORM) Glass platters.

Optical has much faster positioning and for WORM Glass platters have a much higher life expectancy than tape. MO and DVD/CD-R, depending on the external conditions, could last longer than some tapes. Optical media's main advantage is access time compared with tape. The fastest tape can position in about 10 seconds after being loaded while optical is ~25 msecs.

Tradeoffs

Things get dicey as we are in a religious area for many vendors and users, so I will try and accurately present the tradeoffs.

Until the introduction of LTO Ultrium, helical densities far exceeded linear densities, which is often why helical tapes were being used. LTO Ultrium might be changing this. Often helical tape drives can write faster than linear. The Sony DTF-2 drive is a good example where the uncompressed write speed is over 22 MB/s.

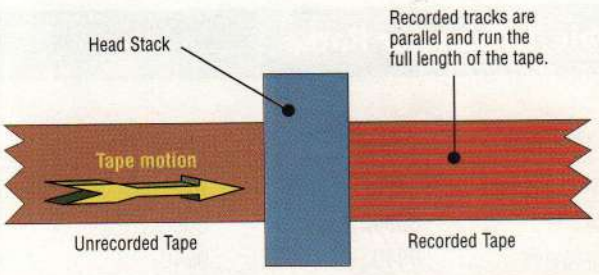
One of the drawbacks of helical with high-density recording is the difficulty it introduces in error detection and correction. A very small defect on a helical tape will probably corrupt the data if the error correction buffer is full. Error correction space is often left on the tape, and if that space fills up the tape becomes unreadable.

In general, helical tape heads wear out long before linear tape heads, given that the tape heads make more intimate contact with the tape.

Reliability is generally higher for linear tapes over helical for both the media and head life of the drive because more contact means more wear.

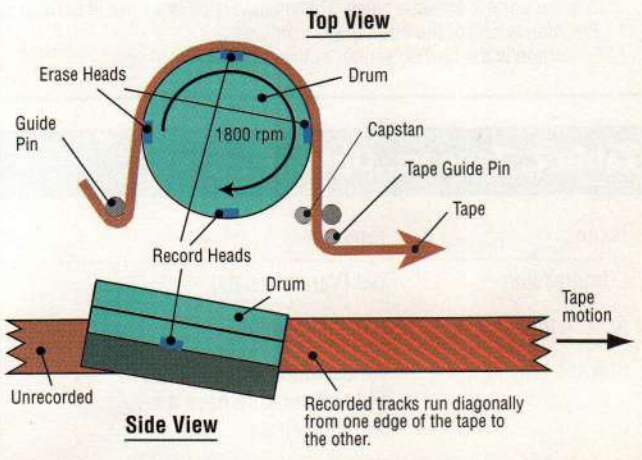
Transfer rate is faster for helical comparing the fastest linear drives vs. the fastest helical drives. Because of media wear, high-end linear tapes generally have a longer storage life than high-end helical tapes.

Figure 1. Record Tracking



Recording tracking for linear tape drives is parallel.

Figure 2. Helical Scan System



Records data in diagonal stripes across the tape.

Optical/Tape Tradeoffs

Optical and tape really should not be compared together as they have such different characteristics (see Table 2).

So what do you do? One other important consideration for tapes and optical storage for long periods is not necessarily the life of the media but the life of the interface. Take the 7-track tapes from the 1970s. These tapes had a supposed shelf life of 20 years, but I am not sure that many people could find 7-track tape drives in the late 1990s to read those tapes. Even if the media lasts for 500 years, the interface and drive might last no longer than 10 years. Storage migration plans and concepts must be developed and implemented.

If you need fast access and highly reliable storage under harsher conditions of high temperatures over 78 F and humidity over 60% RH go with optical. The cost per GB will be much higher than tape. Tapes have compression, but that compression varies significantly. You must test compression with your own data. I have seen on the same data drives compress 1-to-1, 2-to-1, 4-to-1, and 6-to-1 with different data set. Think of the media costs!

You have a large number of variables to consider and factor into your secondary storage media decision. The data reliability verses cost tradeoffs for optical might suggest that two copies on tape are cheaper than one optical copy. Maybe in five years this will all be a moot point if the promises of holographic storage really do work, but that's another column some month in the future.

Next Month

I was going to cover libraries and jukeboxes this month, but realized that secondary storage was going to take the whole column. So next month we will discuss tape libraries and optical jukeboxes. See you then. Of course if anyone has any column suggestions please let me know. ⊷

Henry Newman *is a 20 year veteran in high-performance systems working mostly in I/O and system tuning. He is currently employed by a consulting company.* hnewman@cpg.com.

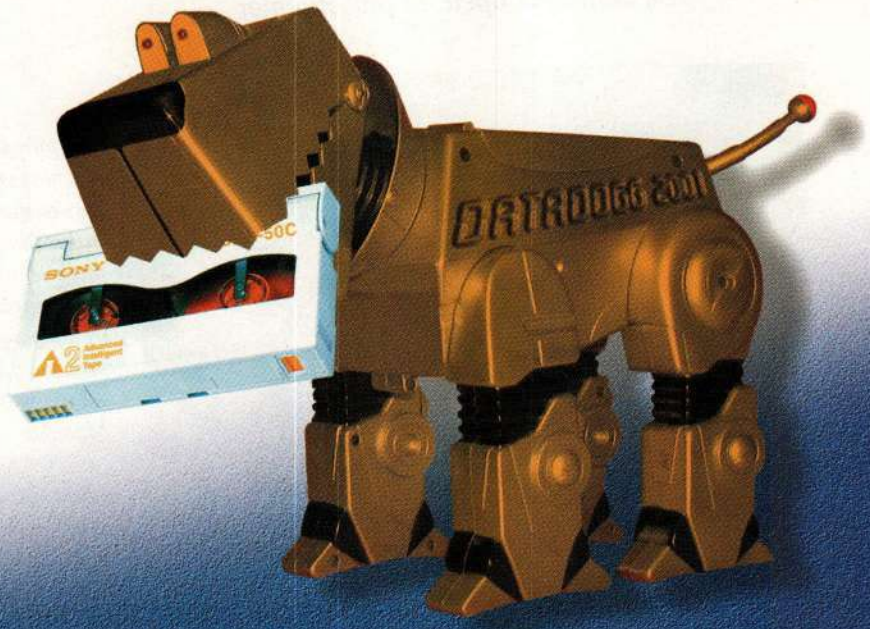
Table 1. Transfer Rates

Vendor/Type	Drive	Media	Year Introduced	Capacity	Peak Transfer Rate MB/s uncompressed
IBM	3420	Reel-to-Reel	1974	150 MB	1.25
IBM	3590E	3590	1999	40 GB	14
Storagetek	9940	9940	2000	60 GB	10
Sony	GY-8240FC	DTF-2	2000	200 GB/60 GB*	24
LTO	Ultrium	Ultrium	2000	100 GB	~14
Optical M/O	MOD910e/i**	P9100E	2000	9.1 GB	6.14
Optical	Plasmon 14 inch	TrueWORM	2000	30 GB	6.0 read/write
WORM	WORM***	8000	-	-	2.9 write verify

* Sony supports 2 cassette types. The small version does not fit in most robots while the large is supported by both ADIC and Sony robots
** Plasmon is one of the leading MO vendors
*** Plasmon is the leading vendor in this technology

Table 2. Tape vs. Optical

Issue	Tape	Optical
Compression	Yes (Varies widely)	None
Access time	10-100 seconds	25 msecs
Storage Life	Varies widely but humidity and temperature have a dramatic effect	Generally 30-100 years with much less impact from humidity and temperature
Security of data	Few vendors have WORM tapes	WORM glass platters are widely accepted as safe

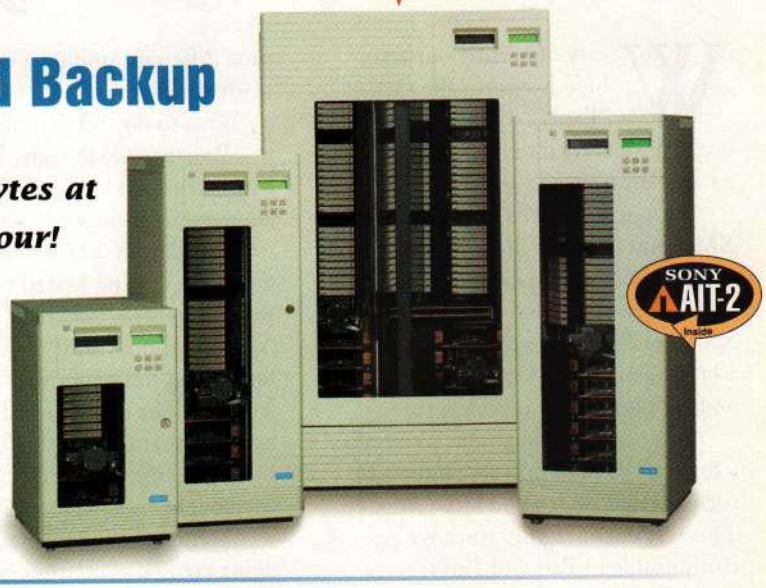


There Are Better Ways To Automate Your Backup:

Sony® AIT Tape Libraries From Qualstar®

Automated Unattended Backup

Stores From 1.5 to Over 78 Terabytes at Speeds Up to 670 Gigabytes Per Hour!



- Very Cost-Effective Data Storage Solutions
- Serverless Backup
- 24 x 7 Onsite Support
- Unsurpassed Warranty
- SANSmart Fibre Channel Connectivity

Call us, or visit our web site today and see what happens when we combine Qualstar Automated, Unattended Backup Solutions with Sony AIT Tape Technology.

1-800-468-0680 • TEL: (805) 583-7744 • FAX: (805) 583-7749 • E-MAIL: sales@qualstar.com • WEB: www.sansmart.com

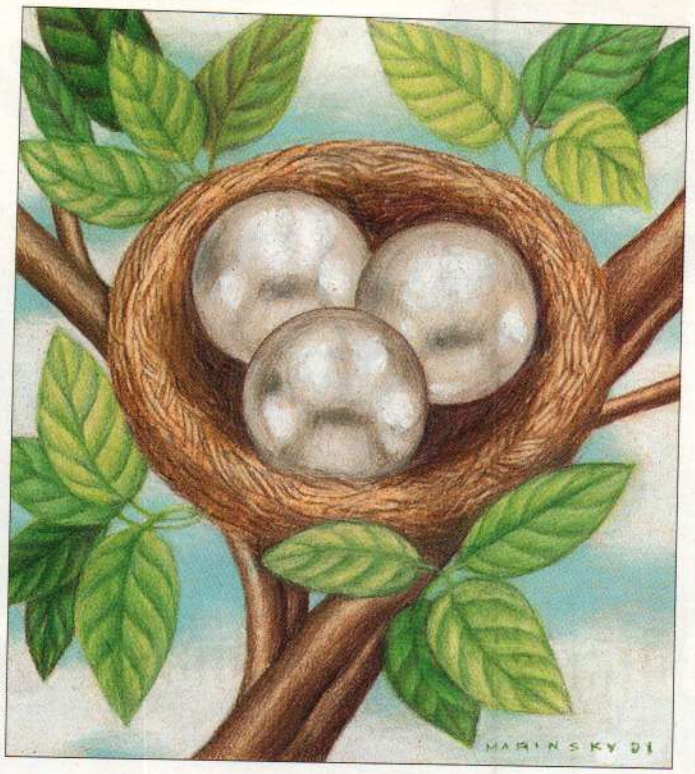
QUALSTAR®

SIMPLY RELIABLE

AIT is a registered trademark of SONY.
Qualstar, the Qualstar logo, SanSmart, and The Tape Library Experts are Registered Trademarks of Qualstar Corporation. NASDAQ: QBAK

The Tape Library Experts™

by Jeffreys Copeland and Haemer



JANE MARINSKY

The purpose of a programming system is to make a computer easy to use. To do this, it furnishes languages and various facilities that are in fact programs invoked and controlled by language features. But these facilities are bought at a price …
– Fred Brooks, *The Mythical Man Month*

There was an Old
Man with a beard,
Who said, "It is
just as I feared!
Two Owls and a Hen,
Four Larks and a Wren,
Have all built their
nests in my beard!"
– Edward Lear,
Book of Nonsense

The Nesting Instinct

Weary of beating ourselves bloody against the rock of "Expect scripts in Perl," we decided we needed to try something easier and less stressful.

Our boss had asked us to look into XML, and writing a little XML-related code seemed like a good place to start. (We'll warn you right off that this isn't an XML column; still, we'll need to tell you a little about XML–a cousin to HTML–and we'll get to that in a second.)

A quick scan through the Web finds a forest of Perl modules to handle XML, such as `XML::Twig` and `XML::Grove`. This was encouraging, since we are comfortable in Perl and Perl is terse and powerful: we can write interesting Perl programs in few enough lines to fit in a 2,000-word column.

But wait. There are about 50 XML modules on the CPAN. Which one shall we use? O'Reilly and Associates' `xml.com` has an article, "Ways to Rome: Processing XML with Perl" that contrasts 11 different approaches to XML using nine different modules. We were not ready for that.

What to do? What to do?

"Perhaps," we thought, "a fresh look at things would help." A little free-association got us an answer.

One view of XML is that it's a language for writing nested text. Any time we hear "nested," we think "parsers" and "stack machines" and "pumping lemmas for context-free grammars." (Well, one Jeff does. The other one thinks about the hatching sparrow's egg he found on the sidewalk earlier this week.) "Maybe this," we thought, "is a job for *yacc*."

Careful readers have noticed that we didn't say free association got us a *good* idea. But what of that?

Besides, we had just installed RedHat 7.1 (on special for $19.95 at our local CompUSA), and writing a *yacc* parser for XML on a system where lots of things were no longer where we expected them to be seemed, oh, challenging. To round out the experience, we decided to switch from *vi* to *emacs*.

Thoughtful study of the Greek and Latin classics teaches that even watching people get torn apart by lions can have entertainment value as long as you have good seats.

So relax, don't worry, have a home brew, read on.

XML is a Good Idea

XML looks like this:

```
<?xml version="1.0"
  standalone="yes"?>
<order>
 <customer>
   <name>Coyote, Ltd.</name>
   <shipping_info>
 <address>1313 Desert
Road</address>
 <city>Flagstaff</city>
 <state>AZ</state>
 <zip>90210</zip>
   </shipping_info>
 </customer>
 <item>
   <product id="1111">Acme
```

```
Rocket Jet Pack</product>
    <quantity type="each">1</quantity>
  </item>
  <item>
    <product id="2222">Roadrunner Chow</product>
    <quantity type="bag">10</quantity>
  </item>
</order>
```

It's data dressed up to look like HTML.

Notice a couple of things. For one, it's structured, but familiar. Orders have a customer followed by items. Customers have names and shipping information. Shipping information has a street address, city, state, and zip. And so on.

The syntax for beginning and ending tags that everyone is now used to from HTML, `<some_tag>` and `</some_tag>`, are borrowed and applied to any tag you want to make up.

Actually, both HTML and XML steal this syntax from their parent, SGML (Standard Generalized Mark-up Language), which gave everyone a way to mark up text that was so generalized that no one would use it.

HTML provided a very simplified subset of SGML and proved that something SGML-like could succeed. But if you've ever wanted to define your own tags in HTML–and, lets face it, we all have–you'll know that HTML can sometimes be too simplified. Enter XML, a simplified version of SGML that still lets you define your own tags.

Second, it's text, and there's a public spec. There's no shortage of structured document formats. Microsoft Word documents are structured in some way. So are Excel spreadsheets.

The next time you try to write a shell script to parse one of these to extract and reformat information needed for another document, drop us a line to tell us how much fun you're having. We need the laugh.

XML files, in contrast, are human-readable. For documents alone, this is a great idea, but it's an idea in an arena where there are already a lot of mature, open, text-based markup languages, such as *troff*, TeX, and HTML.

What transforms XML into something interesting is turning simple text markup on its head by realizing that XML tags can be used to attach meanings to any kind of nested data.

For example?

Well, spreadsheets, which we mentioned earlier, are one. The Gnome spreadsheet, *gnumeric*, uses XML as its native-file format. Run *gnumeric*, type a few things in, save the spreadsheet, then gunzip it (the files are gzipped to save space) and look at it using whatever you use to peruse ASCII files: *cat*, *less*, *vi*, whatever.

Glade, Gnome's user interface builder, spits out interface description files in XML.

For that matter, any data structures you want to pass around can be flattened out with XML, written to a file, and then read in by anything else that understands the same tags. SOAP and XML-RPC, for example, are RPC (remote-procedure-call) mechanisms that pass arguments and return values over the Web as XML-formatted data. WebDAV (Distributed Authoring and Versioning)–Wiki-Wiki Web filtered through too many committees, sort of–passes its data around as XML, too.

And, no matter how intricately structured, no matter how complex the semantics, XML is always something you can read. With *cat*, with *less*, with *vi*. (OK, even with *emacs*.) You can hunt for stuff in it with *grep*. You can print it out on your printer.

It's just text.

Validating XML

What else can we do with XML? We can parse it.

Why is that interesting? Because syntactically valid data, like syntactically valid programs, are a long way towards being correct.

Dimidium facti, qui coepit, habet. –Horace

Well begun is half done.

We find that using `perl -c -w -Mstrict` on our Perl programs goes a long way towards making them work.

When developing data formats as complex as the output of *gnumeric*, having a tool verify that data have the correct structure is also a big help.

Moreover, when data are known to be valid, you don't have to scatter data validation through your programs. No need to write *N* separate sets of routines (with *N* separate sets of bugs) for *N* separate application, to check that each customer has a city, state, and zip code if there is a central data-validator or a single data-validation library that anyone can call.

There are a lot of articles on XML validation, but we particularly like Kip Hampton's "Simple XML Validation in Perl," (`http://www.xml.com/pub/a/2000/11/08/perl`), because it uses Joshua Nathaniel Pritikin's `Test.pm`, which we wrote about a few months back, ("The Art of SoftwareTesting," `http://swexpert.com/C9/SE.C9.AUG.00.pdf`, and "Testy, Aren't We?", `http://swexpert.com/C9/SE.C9.SEP.00.pdf`).

Better still, it uses the module in a way we suspect its author never intended: as part of a stand-alone application.

Rolling Our Own

As we mentioned at the outset, we wanted to see if we understood the concepts, so we decided to build our own XML validator. Since we were impressed by Hampton's paper, we started with his data. These are the customer-order data we showed earlier, available for downloading from O'Reilly's `xml.com` Web site.

And, so as to not ape his approach (or that of any sane person), we thought it might be fun to validate XML the same way we validate Perl: with a parser. To build our parser, we used the universally-available parser-generator, *yacc*.

Before we show you what we ended up with, let's digress for a minute to review some basic Computer Science.

Lack and Yecchs

A compiler's job is to take a program we feed it, tease apart the text to see what it's supposed to do, and then generate machine instructions that correspond to our source.

The "tease apart" section is done, conceptually, in two steps: tokenizing and parsing.

The first, also called "lexical analysis," just breaks the source up into words. In the code fragment

```
if (year==2001) { riley=16 }
```

the tokens are `if`, `(`, `year`, `==`, `2001`, `)`, `{`, `riley`, `=`, `16`, and `}`. Lexical analysis actually can be done by the program

```
perl -pe 's/$RE/\n/'
```

where `RE` is some tedious, incredibly complex, Perl regular expression. You can get from the grammar of the language to the correct Perl regular expression by buying enough beer for some nearby Automata Theorist.

The second step, parsing, verifies that the tokens are in the right order. In the code fragment

```
} { year riley if = == 2001 16 ) (
```

the tokens are the same, but they aren't arranged into syntactically valid C. This step actually can't be done by any tedious Perl regular expression, no matter how complex it is or how much beer you're willing to buy.

The distinction here is that regular expressions are a way to describe, in a single string, a kind of "machine" called a *finite automaton*, and you can prove that finite automata are not able to do paren matching, or brace matching, or anything that requires matching the beginnings of nested structures with their ends. Such jobs are handled by more powerful machines called *parsers*. Parsers can also tokenize, but that's using a hand grenade for a hand-ax.

Early on, all tokenizers and parsers were hand-crafted, but that was tedious. Before long, at Bell Labs, Mike Lesk and E. Schmidt wrote *lex*, a program that writes tokenizers, and Steve Johnson wrote *yacc*, "yet another compiler-compiler," which writes parsers.

Lex takes an input file full of regular expressions and writes C that compiles into a tokenizer. *Yacc* takes an input file full of rules that look sort of like a BNF grammar and writes C that compiles into a parser. (If you don't know what a BNF grammar is–it stands for "Backus-Naur Form," after its inventors–it'll be obvious from the example below.)

These or their GNU replacements, *flex* and *bison*, have been distributed with UNIX systems since the mid-to-late 1970s.

Rubber and Road

By now, you've undoubtedly noticed that validating an XML file requires two steps: breaking it apart into valid tokens, like `</address>`, and ensuring they're correctly nested and in the proper order. This, then, is what brought *lex* and *yacc* to mind.

Writing *yacc* grammars and *lex* source files is straightforward, so without further ado, here's our *lex* input file.

```
/* $Id: lexer.l,v 1.8 2001/06/02 22:10:45 jsh Exp $ */
%{
#include "y.tab.h"
%}

WS   [ \t\n\r]+
OTHER     [^<>]+

%%

{WS}          /* eat white space */
\<\?xml[^>]*\?> return XML;
\<order> return ORDER;
\<\/order>   return _ORDER;
\<customer>       return CUSTOMER;
\<\/customer>     return _CUSTOMER;
\<name>        return NAME;
\<\/name>          return _NAME;
\<shipping_info>   return SHIPPING_INFO;
\<\/shipping_info>     return _SHIPPING_INFO;
\<address>     return ADDRESS;
\<\/address> return _ADDRESS;
\<item>        return ITEM;
\<\/item>          return _ITEM;
\<city>        return CITY;
\<\/city>          return _CITY;
\<state>       return STATE;
\<\/state>   return _STATE;
\<zip>         return ZIP;
\<\/zip>           return _ZIP;
\<product[^>]*> return PRODUCT;
\<\/product> return _PRODUCT;
\<quantity[^>]*> return QUANTITY;
\<\/quantity>      return _QUANTITY;
<<EOF>>         return 0;
{OTHER}       return STRING;

%%
```

We won't drag you through the lexical dirt, but you can see that the file has a rule for each kind of token that needs to be recognized, including a rule to throw away white space:

```
{WS}              /* eat white space */
```

which refers to an earlier rule that defines white space precisely, as one or more white-space characters (blanks, tabs, newlines, or carriage returns) in a row:

```
WS       [ \t\n\r]+
```

Next, we trot out our yacc input file:

```
/* $Id: parser.y,v 1.9 2001/06/02 22:24:03 jsh Exp $ */

%token ORDER _ORDER CUSTOMER _CUSTOMER NAME _NAME
%token SHIPPING_INFO _SHIPPING_INFO
%token ADDRESS _ADDRESS
%token CITY _CITY STATE _STATE ZIP _ZIP
%token ITEM _ITEM PRODUCT _PRODUCT
```

```
%token QUANTITY _QUANTITY
%token STRING XML

%% /* Grammar rules and actions follow */

input:    /* empty */
  | input xml order
  ;
xml:      XML ;
order:          ORDER customer items _ORDER ;
customer: CUSTOMER name shipping_info _CUSTOMER ;
name:           NAME string _NAME ;
shipping_info: SHIPPING_INFO address city state
zip _SHIPPING_INFO ;
address:  ADDRESS string _ADDRESS ;
city:           CITY string _CITY ;
state:          STATE string _STATE ;
zip:      ZIP string _ZIP ;
items:          item
    | items item
    ;
item:           ITEM product quantity _ITEM
    | ITEM quantity product _ITEM
    ;
product:  PRODUCT string _PRODUCT ;
quantity: QUANTITY string _QUANTITY ;
string:         STRING ;

%%

/* Lexical analyzer returns a double floating point
number on the stack and the token NUM, or the
ASCII character read if not a number.  Skips all
blanks and tabs, returns 0 for EOF. */

#include <stdio.h>
#include <ctype.h>

main ()
{
 yyparse ();
}


#include <stdio.h>

yyerror (s)  /* Called by yyparse on error */
  char *s;
{
 printf ("%s\n", s);
}
```

And, looking back at Kip's code, which do we like better? His.

Reflections

So what did we learn from this exercise?

Well, first, figuring out how to bend *lex* and *yacc* to our will was a lot of work. Unlike Perl, *lex* and *yacc* are tools that we rarely use. But, as with Perl, wielding them efficiently requires practice.

Perl also has modules, syntax-checking, and a debugger, all of which ease development. If we were writing a lot of parsers, or one really big one, we'd probably use *yacc*, but we'd write tools to make our work easier. Normally, each grammar rule would have an action associated with it, so the parser could generate code. We would build C++ classes so the generation of code from disparate actions could be insulated.

Second, we think that Kip's Perl example is clearer and easier to modify. Some of this is for reasons we already mentioned, but some of it is that the modules he uses are tailored to the job of parsing XML. Like SGML, *lex* and *yacc* are general tools, with more power than this job needs. When we see that much power walking down the road, verbosity and syntactic complexity are often walking on either side.

Third, we are impressed, in retrospect, by a few tools that seem to have clean-burning, easy-to-use power. Regular expressions are a great example. To underscore this, can anyone suggest a clean, simple syntax that would let a Perl programmer specify context-free grammars as easily as we can specify regexes?

Fourth, we offer a tiny piece of psychological insight: to wit, when we finished writing our parser and lexer, we knew we were done because feeding our test case to the validator just gave us back a prompt and nothing else. This is, we discovered, mildly depressing. Wedded for hours to this little project, in the end we'd brought together something old (*yacc* and *lex*), something new (XML) and something borrowed (Kip's example), and wound up blue.

Fortunately, there's plenty of fun to be had around every corner, and we're certain you'll spend the next month having some of it.

But do we now feel more comfortable with XML? Well … yeah, actually, we do.

So, until next time, happy trails.

PS: This column was written on June 2, 2001. On this occasion, the Jeffs wish JSH's alienated first-born, Riley Sherman Haemer, a happy birthday. Though she is lost forever to the Boulder, CO, divorce courts, we both love her and will forever miss her more than words or tears will tell.

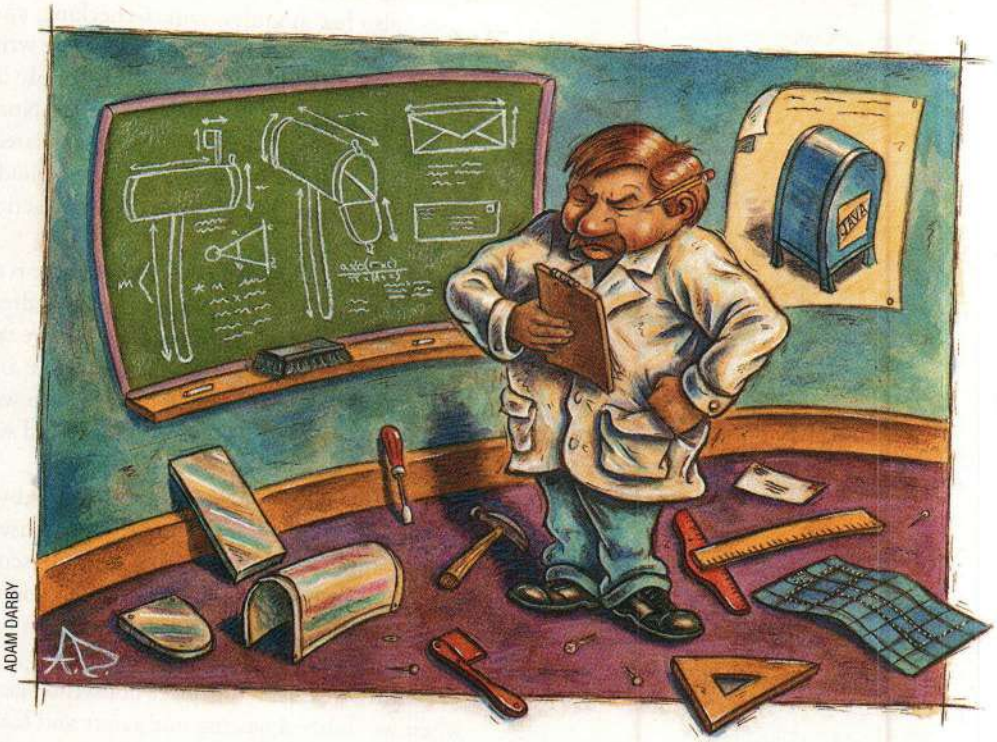
Happy trails, little Riley. ➸

Jeffrey Copeland *(copeland@alumni.caltech.edu) is currently living in the Pacific Northwest, where he spends his time writing UNIX software in a large development organization and fighting damp rot.*

Jeffrey S. Haemer *(jsh@usenix.org) works at Minolta-QMS Inc. in Boulder, CO, building laser printer firmware. Before he worked for QMS, he operated his own consulting firm and did a lot of other things, much like everyone else in the software industry.*

Note: The software from this and past Work columns is available at `http://alumni.caltech.edu/~copeland/work` *or alternately at* `ftp://ftp.cpg.com/pub/Work`.

by Jim Frost



ADAM DARBY

Postal: Design Considerations

In our ongoing quest to make a production-quality IMAP e-mail client in Java, this month we're going to take steps to improve its reliability and hopefully form the foundation for later performance and feature enhancements.

Postal is too large to include in its entirety in print, so you are strongly encouraged to download the package, which includes both source code and the compiled application, from `ftp://ftp.cpg.com/pub/JavaClass/08.2001/postal.tar`.

We have come to an important point in the development of Postal. We have some significant problems we need to fix that affect a large fraction of the code base. We want to fix the problems, but do it in such a way as to avoid rewriting a lot of code. What's more, we have limited resources to apply to the problem–I still have my day job, so Postal is written on nights and weekends–time I could also spend playing with my daughter or riding my mountain bike.

The basic problem is the JavaMail library on which Postal is based has no built-in error recovery features; even simple failures can invalidate quite a lot of objects. There are two primary types of failures: The loss of the connection with the IMAP server, or loss of control of a folder (usually as a result of some other client opening it). Either case can invalidate a number of JavaMail-generated objects, but the first case is devastating: All `Folder` and `Message` objects obtained from that server are invalid as soon as that connection is lost. Everything with a reference to these objects needs to get new ones after the connection is recovered–and that's a lot of stuff in Postal.

Since JavaMail is designed as a set of interfaces, the theory is you can write a new implementation and plug it in to the library or substitute new implementations for the JavaMail implementations wherever you like. If we so desire we can produce a set of `Store`, `Folder`, `Message` and other associated objects that encapsulate the JavaMail objects and, in case of failure, recreate as many of them as necessary behind the scenes to restore functionality.

This approach is particularly appealing because it would mean very localized changes to the bulk of Postal; most of the code would never know the implementation has changed, it would just appear that the objects failed less often. As fine as this theory is, in practice it turned out to be extremely difficult. JavaMail objects are quite complicated to implement; after several hours of trying to get some basic objects implemented I decided it was more work than it was worth.

The alternative is to create our own set of interfaces, exposing only the functionality we actually use in Postal, and build the same kind of recovery system into the implementation. The upside to this is that the interfaces will be much simpler than those of JavaMail, and the implementations will be correspondingly simpler too. The downside is that we'll have to go around and look at all the

code that touches JavaMail objects and modify it to use our interfaces. A quick look at the code found more than 1,200 references to the `Store`, `Folder`, and `Message` types and many more method references. It's quite a lot to change.

Luckily, there is middle ground, in two respects.

First, if we can modify only *some* references we may be able to get a lot of the benefits without huge code changes. We can then incrementally introduce the new interfaces to the remainder of the code base over time, gradually phasing out all use of the JavaMail objects outside of our own interface implementations.

Second, if we can keep many of the method signatures the same between the original JavaMail interfaces and our new versions, we won't have to visit every method call–only the variable declarations. But there's a big downside: Some of our code will be using our new interfaces and some of it the old interfaces, making the code more difficult to understand.

This is a classic software engineering compromise: When there's too much to do it right, design something that lets you do most of the job now and leaves the door open for cleaning up the rest in a later release.

Listing 1. *JavaMailManager*

```
public class JavaMailManager
    implements ServerListener
{
    // ... code ellided ...

    /** Mailer configuration object. */
    private Configuration configuration;
    /** Set of mail stores by name. */
    private HashMap serversByName = new HashMap(11);
    /** Cache of open folders. */
    private Map foldersByPath = Collections.synchronizedMap(new HashMap(11));
    /** Exclusive access lock, to avoid multithreading issues with the
     * JavaMail library. */
    private ReentrantLock lock = new ReentrantLock();

    public JavaMailManager(Configuration configuration)
    {
        this.configuration = configuration;
        configuration.getDispatcher().addServerListener(this);
        // populate ourself with the initial list of servers
        Iterator serverIter = configuration.listServers();
        if (serverIter != null) {
            while (serverIter.hasNext()) {
                Server server = (Server)serverIter.next();
                serversByName.put(server.getServerName(), server);
            }
        }
    }

    // ... code ellided ...

    /** Returns the folder with the indicated path. */
    public Folder getFolder(String folderPath)
    {
        Throwable failureException = null;
        try {
            lock();
            Folder folder = (Folder)foldersByPath.get(folderPath);
            if (folder != null)
                return folder;
            for (int attempt = 1; attempt <= 2; attempt++) {
                try {
                    StringTokenizer tokenizer = new StringTokenizer(folderPath, "/", false);
                    // first token is the server
                    if (!tokenizer.hasMoreTokens())
                        return null;
                    String serverName = tokenizer.nextToken();
                    Server server = (Server)serversByName.get(serverName);
                    if (server == null) {
                        recover(folderPath, null);
                        return null;
                    }
                    Store store = server.getStore();
                    if (store == null) {
                        recover(folderPath, null);
                        continue;
                    }
                    folder = store.getDefaultFolder();
                    if (folder == null) {
                        recover(folderPath, null);
```

Continued on Page 34

What's New

What I chose to do was to create a `FolderHandle` interface (see source distribution) and substitute it everywhere that we previously used a `Folder`. I started by recoding the `Server` and `FolderSelectorPanel` objects and `FolderListener` interface and worked my way down through the code base by letting the compiler point out the places where types were now mismatched. In an hour or so I'd caught most of them, enough to start plugging in code to see what else broke.

Implementing the `FolderHandle` was more problematic because many of the errors affecting a `Folder` are actually inherited from errors in the `Store` object (represented by the `Server` object in Postal). We may discover them when manipulating the `Folder`, but we need a larger-scope recovery framework to actually fix them.

The top of that framework is the `JavaMailManager` (Listing 1, Page 33), whose job is two-fold: Look up JavaMail `Folder` objects, and recover those objects after failure. By doing this we can have our `FolderHandle` objects refer to `Folder` objects symbolically except when they're actually being used,

```
                    continue;
                }

                // iterate the rest of the path elements to find the folder
                while (tokenizer.hasMoreTokens()) {
                    String folderName = tokenizer.nextToken();
                    folder = folder.getFolder(folderName);
                    if (folder == null)
                        break;
                }
                if (folder != null)
                    foldersByPath.put(folderPath, folder);
                return folder;
            }
            catch (MessagingException e) {
                failureException = e;
            }
        }
    }
    finally {
        unlock();
        if (failureException != null)
            getDispatcher().fireErrorEvent(failureException);
    }
    return null;
}

/** Called if there is an error accessing a folder.  This attempts
 * to recover from the error such that a subsequent  access
 * attempt will succeed. */
public void recover(String folderPath, Throwable exception)
{
    // a folder might get closed if some other client accesses it.  that's
    // not terminal; re-opening it should allow it to work.
    if (exception instanceof FolderClosedException) {
        getDispatcher().fireStatusEvent(this, MessageFormat.format(REOPENING_FOLDER,
                                                                   new String[] { folderPath }));
        Folder folder = getFolder(folderPath);
        try {
            if (folder != null)
                folder.open(Folder.READ_WRITE); // only mode we support right now
            return; // everything should be good now
        }
        catch (MessagingException e) {
            // something's badly broken, let's invalidate the whole server
        }
    }
    // if we get this far we aren't sure what the problem is, so we try
    // to invalidate all folders on the server and reconnect.
    StringTokenizer tokenizer = new StringTokenizer(folderPath, "/", false);
    if (!tokenizer.hasMoreTokens())
        return; // huh?
    String serverName = tokenizer.nextToken();
    getDispatcher().fireStatusEvent(this, MessageFormat.format(RECONNECTING_SERVER,
                                                               new String[] { serverName}));
    Server server = (Server)serversByName.get(serverName);
    if (server == null) // somehow the server disappeared
        return;
    server.close();
    StringserverPrefix = "/" + server.getServerName() + "/";
    ArrayList invalidFolders = new ArrayList(10);
    // collect list of folders that are found on this server
```

Continued on Page 35

dramatically limiting the impact of a failure. Notice that `JavaMailManager` also implements a `Folder` cache; once it looks up a `Folder` object it remembers it until something tells it to forget. This is important since our code will be looking up `Folder` objects more often than it did in the past.

`JavaMailFolderHandle` (Listing 2) implements the fault-tolerant `FolderHandle`. Almost all of the work is done by `syncMessages()`, which requests the message list from the `Folder` and builds a cache of messages similar to the folder cache in `JavaMailManager`. Notice the loop surrounding the try block and most of the method body. This is our failure recovery loop. If we get a failure trying to access the folder's messages we call `recover()` (which simply calls the all-encompassing `JavaMailManager.recover()`) to clear out the `Folder` cache, then we try again. If all goes well, the recovery will fix the problem and we'll get a working `Folder` the next time around. If all *doesn't* go well we'll fail the second time too, at which point we give up and tell the user that something is wrong. We could keep trying, but usually if the error cannot be recovered in one try it's pretty serious and we'll just be wasting our time.

```
        Iterator folderIter = foldersByPath.keySet().iterator();
        while (folderIter.hasNext()) {
            folderPath = (String)folderIter.next();
            if (folderPath.startsWith(serverPrefix))
                invalidFolders.add(folderPath);
        }
        // remove them all
        folderIter = invalidFolders.iterator();
        while (folderIter.hasNext()) {
            foldersByPath.remove(folderIter.next());
        }
        server.connect();
    }

    /** Locks the manager from concurrent accesses. */
    public void lock()
    {
        for (;;) {
            try {
                lock.acquire();
                return;
            }
            catch (InterruptedException e) {
                continue;
            }
        }
    }

    /** Unlocks the manager from concurrent accesses. */
    public void unlock()
    {
        lock.release();
    }

    // ... code ellided ...
}
```

Listing 2. *JavaMailFolderHandle*

```
public class JavaMailFolderHandle implements FolderHandle
{
    // ... code ellided ...

    /** Mail manager used to look up the JavaMail folder. */
    private JavaMailManager mailManager;
    /** Name of the folder. */
    private String name;
    /** Full path to the folder. */
    private String path;
    /** Parent of this folder. */
    private FolderHandle parent;
    /** List of child folder names, or null if there aren't any. */
    private String[] childFolders;
    /** Map of child folders indexed by folder name. */
    private Map childFoldersByName = Collections.synchronizedMap(new HashMap(11));
    /** List of messages IDs in the folder. */
    private String[] messageIds;
    /** Map of messages indexed by message ID. */
```

Continued on Page 36

Observant readers will have noticed that `JavaMailFolderHandle` doesn't cache `Message` objects, but rather `JavaMailMessageHandle` objects. We don't have the time this month to re-write all of the code that deals with `Message` objects, but we've recoded the `MessageSelectorModel` to use `MessageHandles` in preparation for that migration. This also improves our reliability significantly since the `MessageSelectorModel` touches more messages than anything else in the application.

While we were re-writing `MessageSelectorModel` we made one other significant change: We eliminated the refresh thread, moving it into the `JavaMailFolderHandle`. Now whenever the folder is refreshed it sends a change notification and the `MessageSelectorModel` picks up the changes. This reduces the complexity of the `MessageSelectorModel` and makes it possible for us to refresh multiple folders in parallel for better interactivity. We're not making use of that yet, but we will.

Managing Concurrency

One of the more serious issues Postal has had with JavaMail is its non-reentrancy. If we attempt to perform two operations

```
private HashMap messagesById = new HashMap(89);
/** If true, child folder information is assumed to be up-to-date. */
private boolean validFolders;
/** If true, message information is assumed to be up-to-date. */
private boolean validMessages;

public JavaMailFolderHandle(JavaMailManager mailManager,
                            FolderHandle parent,
                            String name)
{
    this.mailManager = mailManager;
    this.parent = parent;
    this.name = name;
    this.path = parent.getPath() + "/" + name;
}

// ... code ellided ...

/** Called whenever an exception occurs during folder access in
 * order to attempt failure recovery for a re-try. */
private void recover(Throwable t)
{ mailManager.recover(path, t); }

// ... code ellided ...

/** Reads messages from server if necessary. */
private void syncMessages()
{
    if (validMessages)
        return;
    RefreshOperationStatus refreshStatus = new RefreshOperationStatus(this, REFRESH_OPERATION,
                                                                      getDispatcher());
    refreshStatus.begin(REFRESHING_MESSAGE_LIST);
    Throwable failureException = null;
    try {
        lock();
        for (int attempt = 1; attempt <= 2; attempt++) {
            try {
                failureException = null;
                synchronized (this) {
                    messageIds = null;
                }
                messagesById.clear();
                Folder folder = mailManager.getFolder(path);
                if (folder == null) {
                    // complain about missing folder here
                    continue;
                }
                // if folder can't contain messages, we're done
                if ((folder.getType() & Folder.HOLDS_MESSAGES) == 0)
                    return;
                if (!folder.isOpen())
                    folder.open(Folder.READ_WRITE);
                Message[] messages = folder.getMessages();
                if (messages == null)
                    return;
                String[] newMessageIds = new String[messages.length];
                for (int i = 0; i < messages.length; i++) {
                    Message message = (Message)messages[i];
                    String messageId = MessageUtils.getMessageId(message, i);
                    messagesById.put(messageId, message);
```

Continued on Page 37

on the same server at the same time, JavaMail fails. A couple of months ago I quietly put in some simple locking code to try to reduce concurrent use of JavaMail, with only limited success. Because we would like Postal to make a lot of use of concurrency, one of the design requirements of the new interfaces is that they must be reentrant. In this month's implementation we serialize all contact with the JavaMail library through our interfaces with the `JavaMailManager.lock()` method, which utilizes Doug Lea's excellent `util.concurrent` library from his book *Concurrent Programming in Java: Design Principles and Patterns,* (ISBN 0-201-31009-0, Addison-Wesley).

We're overdoing it, serializing even across servers, but we can fix that when we finish migrating our `Server` object to the new interfaces.

So how well does all of this work? Pretty well; many errors are now recovered from without user intervention.

That's all we have time for this month. Next month we'll introduce a new feature–printer support. Stay tuned, printing in Java is an adventure you won't want to miss. ⊷

***Jim Frost** is a software engineer specializing in Java technologies and strong opinions.* `jimf@frostbytes.com`.

```
                    newMessageIds[i] = messageId;
                }
                synchronized (this) {
                    messageIds = newMessageIds;
                }
            }
            catch (Throwable e) {
                failureException = e;
                recover(e);
            }
        }
    }
    catch (Throwable e) {
        failureException = e;
        recover(e);
    }
    finally {
        unlock();
        if (failureException != null)
            fireErrorEvent(REFRESH_FAILED, failureException);
        refreshStatus.end();
    }
    validMessages = true;
    fireFolderChangedEvent();
}

/** Requests exclusive access to this folder. */
public void lock() { mailManager.lock(); }

/** Releases exclusive access to this folder. */
public void unlock() { mailManager.unlock(); }

// ... code ellided ...

/** Returns handles to the messages in the folder. */
public MessageHandle[] getMessages()
{
    syncMessages();
    try {
        lock();
        String[] ids;
        synchronized (this) {
            ids = messageIds;
        }
        if (ids == null)
            return null;
        MessageHandle[] handles = new MessageHandle[ids.length];
        for (int i = 0; i < ids.length; i++) {
            handles[i] = new JavaMailMessageHandle(this, ids[i]);
        }
        return handles;
    }
    finally {
        unlock();
    }
}

// ... code ellided ...
}
```

by Cameron Laird

ROBERT BURGER

Fun with SOAP

Let's program SOAP. I think you'll like it. I understand software far better when I get my own hands dirty (or, in this case, SOAPy). My imagination seems to be the limited kind that requires working programs before I can "get" most new technologies. That's one of the reasons I slip executable source into as many "Open Sources" columns as I can justify. I was able to come up with the descriptions of Web services (WS) in June's column, "Defining Web Services," (`http://swexpert.com/CC/SE.C12.JUN.01.pdf`) only after I'd done my own WS experiments. This month, you get the same opportunity. Also, I'll add a footnote that briefly updates April's "Oracle's Open-Source Fast Lane," (`http://swexpert.com/CC/SE.C12.APR.01.pdf`).

SOAP's Business and Technology

Don't let SOAP's media coverage discourage you. Yes, there are several confusing business angles to SOAP. IBM, Oracle, Sun, and Microsoft are just a few of the companies that occasionally talk about how SOAP involves revolutionary changes in business relations, but you can put all that aside. Think of SOAP as just another technology for distributed computing, like CORBA, DCOM, Jini, and dozens of others. These are all ways to exploit multiple networked hosts to solve problems. SOAP simply has a few crucial engineering advantages that make it more fun than any of the others.

This month's examples are all presented in Perl. One of SOAP's attractions is its language-neutrality: in principle, there are bindings for almost any language you might want to use. Even for languages such as Eiffel and Tcl, with rather "unripe" interfaces, progress seems to be rapid. With all these choices, why select Perl? As often as my own coding is in C and Java, I urge use of high-level "scripting" languages whenever possible.

Perl has great SOAP libraries, along with the combination of accessibility and succinctness users deserve. Even if you never use Perl yourself, expect to get plenty from the code examples that you can apply in your chosen language(s).

Assemble the Ingredients

Warm up for this exercise by installing the needed Perl modules. Use version 5.6.0 of Perl or later; making an earlier version work with XML isn't worth the trouble involved. Install the CPAN and `SOAP::Lite` modules.

SOAP programming does not strictly require CPAN. However, the CPAN module simplifies serious Perl work, including installation of the modules that SOAP *does* require. (`http://www.cpan.org/modules/by-module/CPAN/CPAN-1.59.tar.gz` is a handy release). Once you have downloaded and unpacked this, just go through the standard Perl sequence of:

```
perl Makefile.PL
make
su
make install
```

Start an interactive CPAN session as the root user with

```
perl -MCPAN -e shell
```

After a bit of largely automated configuration, request `install SOAP::Lite` at the CPAN application's command prompt. The application should take over and install all `SOAP::Lite`'s prerequisite modules, including URI, libwww, XML-Parse, and IO-stringy.

Now you can program. Try the code in Listing 1.

If you run this script through Perl, you'll see a tiny table with such values as

```
48001:  63.0
48002:  62.0
48003:  62.0
48004:  65.0
48005:  63.0
48006:  61.0
```

You've just built a small application that reports on the current reported temperature for a small area of Michigan, tabulated by zip code.

So What?

Why is this interesting? There's frankly nothing new in it. Remote automations have been built almost as long as there's been an Internet; e-mail servers relied on UUCP transport in the mid-1970s to give information on soft drink dispensers, system load, conference schedules, and other computing essentials. Twenty years ago, I was doing nothing I thought original when I demonstrated security calculations exposed as Telex (!) callback services. CORBA and DCOM are nowadays in continuous mission-critical use at financial and engineering centers, co-operating in remarkably sophisticated distributed computations.

What does SOAP bring to the party? Just convenience. I've lately thought of SOAP as the rough equivalent of automatic transmissions for fleet vehicles. In principle an organization could train all its drivers to use standard transmissions well, and, yes, a manual clutch can have a superior engineering efficiency. None of that particularly matters. Automatic transmissions simply eliminate an entire layer of procedural issues. They're the right choice.

Analogously, SOAP hits an engineering "sweet spot," and makes development of distributed applications qualitatively easier.

Analogously, SOAP hits an engineering "sweet spot," and makes development of distributed applications qualitatively easier.

June's column explained that SOAP encodes all data as XML messages. As hilariously inefficient of bandwidth as that seems to old-time RPC or CORBA programmers, it's the right decision. Standardization of XML solves essentially all current representation problems; anything expressible in Unicode can travel by way of SOAP. SOAP also decouples language decisions from service definition, and that settles another entire class of managerial controversies. CORBA is comparably "neutral"; in principle, it is able to work over the same range of platforms and applications as SOAP. One more SOAP advantage, though, is its exploitation of current transport practices. SOAP's configurability means that it's easy to tunnel SOAP through conventional firewalls and proxies. Doing the same with CORBA remains a matter for experts.

So, although there's no absolute divide between the capabilities of SOAP and other technologies, the packaged bundle is irresistible.

Listing 1. SOAP

```
use IO::Handle;
  STDOUT->autoflush; # Print each line without delay.

  use SOAP::Lite;
  my $client = SOAP::Lite->uri('urn:xmethods-Temperature');
  $client->proxy('http://services.xmethods.net:80/soap/servlet/rpcrouter');
  $starting = 48001;   # Pearl Beach, Michigan.
  for ($zip = $starting; $zip <= $starting + 5; $zip++)  {
    my $response = $client->getTemp(SOAP::Data->type( string => $zip));
    if( $response->fault ) { die $response->faultstring; }
    print $zip, ":  ", $response->result, "\n";
  }
```

Introspection and Social Practice

Perhaps SOAP's biggest advantage is more social than technical. CORBA's heritage is in application development. Client coders work from IDL documents that purport to specify interfaces precisely.

SOAP is looser. The data riches available on the Web to human readers inspire SOAP zealots to want the same information available for computation. There are already plenty of weather services on the Web that allow lookup by zip code; the SOAP used in Listing 1 just wraps one of those up ready for digestion by a computer application, rather than a person.

How will the data be used? Typical CORBA development starts with a statement of requirements, and defines a service to meet them. SOAP appears to be on a more "emergent" developmental path: providers open up SOAP services, then find out how clients are using them.

SOAP is not particularly introspective. A future column which demonstrates WSDL will show how to interrogate a Web service for the methods and data it offers. In practice, though, SOAP's transport of all messages within human-readable XML means that SOAP is generally more "self-documenting" than comparable CORBA applications. When I first wrote the application in Listing 1, I had the method invocation as

```
my $response = $client->getTemp($zip);
```

The result was

```
services.temperature.TempService.getTemp(int) --
no signature match at sample.pl line 15.
```

From there, it was a quick step to make the correction

```
my $response = $client->
getTemp/(SOAP::Data->type(string=>$zip));
```

which tells the SOAP server to regard "48001" as a string, rather than an integer.

That was an amateurish mistake for me to make, because the `xmethods-Temperature` is accurately documented. I should have read the manual. On the other hand, we all know how often documentation is not current, and what a practical benefit good exception messages are. Even with fragmentary information, good programmers can quickly pull SOAP pieces together.

And that's what we're after. It's easy to imagine complete useful applications scripted in SOAP methods. A personal weatherman might recognize the name of a town as you speak it, look up its location, retrieve the weather conditions, and render them back into speech, all by gluing together public SOAP services. The programmer would never have to learn about XML, HTTP, or hardware vagaries; everything just plugs together.

Book Recommendation

My April column made the point that Oracle is one of the largest beneficiaries and benefactors of Open Source. Only after the article was written did I learn about a book which expanded the same topic to more than 400 pages. Andy Duncan and Sean Hull's *Oracle & Open Source* (O'Reilly and Associates, Sebastopol, CA) is a good one I'm happy to recommend. The target audience is working programmers. While there's valuable and correct introductory material on strategic or business aspects of open source, 90% of the text is meaty code, installation tips, definitions, and related references. Duncan and Hull are Oracle specialists whose method is clear and commendable: they've written the book they wish they could have bought five years earlier in their own careers.

Finding and Creating Services

How do you take your own programming beyond Listing 1? Start on the client-side, as Listing 1 illustrates, but with different services. XMethods (`http://www.xmethods.net`) indexes score scores of available free services, along with documentation and sample code that accesses them. Some of these services are quite substantial; they make significant applications possible and even easy. At the time I write this, XMethods lists services to deliver Usenet news, store XML documents off-site, maintain mailing lists, search government securities records, translate between different human languages, and much more.

Once you're comfortable with client-side SOAP work, you'll likely want to offer your own SOAP services. Programming SOAP on the server side with higher-level languages is nearly as easy as client work. Server-side configuration is more delicate, though, because it demands respect for Web service and platform compatibility. Still, it doesn't take long to describe a minimal example that applies directly to most UNIX CGI-enabled servers: create `cgi-bin/Mytime.pm` with the following content:

```
package Mytime;
sub current hour
{
   ($sec, $min, $hour) = localtime(time);
    return $hour;
}
1;
```

Notice the "1" at the end of the source. This is a standard idiom for construction of a properly-formed module. A module must return a value of one to signal that it has been imported correctly; a constant terminal statement with a bare "1" is generally the right way to do so.

Test it by running:

```
use Mytime;
print "The current hour is ",
Mytime::current_hour(), ".\n";
```

On the server machine, in the CGI directory, create a service such as `cgi-bin/example.1`:

```
use SOAP::Transport::HTTP;
use Mytime;

SOAP::Transport::HTTP::CGI
   -> dispatch_to('Mytime')
   -> handle;
```

Now, from a client machine, update Listing 1 to:

```
#!/usr/local/bin/perl5.6.1

use SOAP::Lite;
print "The current hour on the server machine is ",
   SOAP::Lite->uri('Mytime')->
     proxy('http://bread.ecunet.org/cgi-bin/
     example.1')->
     current_hour()->
     result,
   ".\n";
```

Execute the update and you should see `The current hour on the server machine is five`, or perhaps something later. Notice you're programming remote procedure calls without having to learn anything about platform-specific data representations, XML, dispatch description languages, or the other details usually involved in distributed computing.

Summary

So how does SOAP look to a programmer? You can work in any language; there's an existing and growing infrastructure of accessible services; the programming is easy; and the technology is one that has the blessing of all the largest vendors. Though SOAP is a big subject, with an immense array of options and configurability, all it takes to make your own start are the short listings above. Enjoy!

When you do become successful with SOAP, special thanks should go to Paul Kulchenko, author of `SOAP::Lite` and many related works; Tony Hong and James Hong, for their invaluable XMethods Web Services directory, SOAP testing, and related services; and Wales-based "Internet development company" Lucin PLC, which both supplies numerous free Web services, and has open-sourced much of its own SOAP programming.

For more detail on SOAP aspects related to this column, see `http://starbase.neosoft.com/~claird/comp.infosystems.www.misc/SOAP.html`. ➵

Cameron Laird *is vice president of Phaseit Inc.* `www.phaseit.net/`. *He's been extracting results from remote computers since the late '60s.*

SERVER/WORKSTATION EXPERT

EDITORIAL FOCUS

SW EXPERT

Server-Based Computing Solutions

September Datawarehousing for the Enterprise
- Customer Relationship Management (CRM)
- Enterprise Resource Planning (ERP)
- Enterprise Application Integration (EAI) tools

October Java
- Server side-apps
- Enterprise development tools
- Security

November The Bandwidth Squeeze
- Proxy
- Caching
- Load balancing

Visit **http://swexpert.com** for archived issues

X-MEN

DATABASES EMBRACE XML

Whether as a data transport or native storage scheme, databases, often called the 900-pound gorilla of IT, are making room for XML.

BY ALEXANDRA BARRETT, SENIOR EDITOR

In the three years since the World Wide Web Consortium (W3C) recommended version 1.0 of the eXtensible Markup Language (XML), the technology has not stood still for a minute. "XML is an unusual case," says Paul Grabscheid, vice president of strategic planning for database vendor InterSystems Corp., Cambridge, MA, "in that it gets a whole lot of hype, as well as having real value."

If the early days of XML were marked by organizations using it as a data interchange mechanism–for a business-to-business (B2B) or enterprise application integration (EAI) context–today, XML is increasingly being used to mark up original content. "Getting content into XML in the first place is getting easier–editors are much more robust," says Soumitra Sengupta, chief technology officer at B-Bop Inc., Burlingame, CA, an XML software vendor.

People's acceptance of converting documents to XML has changed dramatically in the past 18 months, some vendors note. "Before, people were tentative about XML," says Sengupta. "Now, they are confident it is the right solution."

Analysts seem to agree with this assessment. "XML is the dominant way of describing content for Web applications," states Rob Perry, senior analyst of the Yankee Group. "It's already there, not just moving in that direction."

But with the surge in XML data, developers are increasingly pressed to answer the question, "how does one

ILLUSTRATION BY CHET PHILLIPS

manage all these XML documents?"

In 1999, the California Board of Equalization (CBE), the state agency that processes sales taxes, embarked on a project where that question took center stage. The CBE decided that it was time to allow its taxpayers–a total of almost 900,000 businesses statewide–to complete their tax forms online, through a Web browser.

At the time, the CBE's data architect, Larry Hanson, made a decision that, ultimately, would brand his organization as a trendsetter: use XML as the primary data format. "If we're going to go ahead and do electronic filing, we should move into the new world," he says.

Should one leverage an existing relational database to store XML, or should one store the XML "natively"?

Working with Nationtax Online Inc., Birmingham, AL, an application service provider, the CBE devised a system whereby taxpayers, via the Nationtax application, submit their completed state tax forms in XML format. Assuming that taxpayers are eligible to pay electronically, and that the submitted XML document is valid, the XML data is then stored in its entirety to a database, specifically Tamino, a native XML database from Software AG, Darmstadt, Germany.

But in no way did Tamino displace the CBE's existing database infrastructure–Adabas, also from Software AG, which has long been responsible for processing the data. "We chose Tamino as a complement to our relational database, a place to store the complete XML document," says Hanson. Eventually, the CBE will use Tamino for more than simply archiving the documents. It will open up Tamino internally to data warehousing-type applications.

The CBE's decision to add another database into their mix illustrates an increasingly common dilemma faced by large organizations: should one leverage an existing relational database to store XML, or should one store the XML "natively" in a new class of hierarchical and object databases designed specifically for that purpose?

That question, it turns out, elicits strong emotions from vendors and developers alike.

Dominant relational database vendors, not surprisingly, stand firm that as long as XML data is the same data that runs business, storing it in a scalable, available, transactional relational database is the appropriate choice.

"There have been object database vendors that have been around for a while who have predicted that XML will give them a new lease on life," says John Magee of Oracle Corp., Redwood Shores, CA, and senior director for Oracle 9i product marketing. "But that just hasn't panned out." The reason? "The XML data they're proposing to store is *valuable* data." What customers want, Magee continues, "is the ability to query the database and get formatted XML out the other side, not to move to a whole new infrastructure."

Certainly, no one can accuse big-name relational database vendors of neglecting XML data storage. The big three vendors in the UNIX space, Oracle with Oracle 8 and 9, IBM Corp., Armonk, NY, with DB2, and Sybase Inc., Emeryville, CA, with Adaptive Server Enterprise (ASE) have all long offered several ways to store XML in databases.

Fundamentally, these vendors have offered two basic approaches to XML data storage. The first, explains Tom Traubitz, senior marketing manager for Sybase's ASE, is to store the entirety of the XML document in the database as a Binary Large Object (BLOB) or Character Large Object (CLOB). This approach is certainly easy on developers, but typically delivers only limited text-based query capabilities. The second approach is to parse the XML document against its DTD (document type definition), and reduce it to a relational table. By making the effort of mapping the XML to a relational table, developers thereby gain access to the database's native index and query functions.

In addition, Sybase and Oracle both allow developers to take a hybrid approach, parsing out select elements, and storing the remainder as BLOBs. Either way, storing XML in a relational database works fine, says the Yankee Group's Perry, "up to a point."

But where things start to go wrong is when XML documents parsed to relational tables change their schema, e.g., new elements are added. "What we learned is that the core weakness of our strategy was the parser," says Sybase's Traubitz, which this June announced several XML enhancements to ASE 12.5. The problem? "XML, like its name implies, is a naturally extensible language, and the thing about extensions is they tend to break parsers."

In other words, if the XML document does not have a fixed schema–and many of them don't–developers may find themselves needing to parse their XML on a fairly regular basis. "If mapping is important to [a developer] because they want to be able to query their [XML] document, then yes, they're going to have to change their mapping," says Pat Selinger, IBM Fellow and director of database integration. "But that's life."

But that, says Chris Parkerson, product manager at eXcelon, Burlington, MA, maker of the eXtensible Information Server, or XIS, "is when DBAs start pulling their hair out."

Indeed, among XML developers, you don't have to go very far to find someone discontented with storing XML in a relational database. "No one should be abandoning relational databases," says Paul Brown, president of FiveSight Technologies, a Chicago, IL-based developer of business process automation systems. "It's just that no one should be using them to store XML data."

Instead, Brown and others increasingly advocate using an emerging class of so-called "native" XML databases, of which Software AG's Tamino is one of the better known, but also includes database products from vendors such as Ipedo Inc., Redwood City, CA; Ixiasoft Inc., Outremont, Quebec; XYZFind Corp., Seattle, WA; and open-source projects such as dbXML and Xhive.

FiveSight's Brown cemented his appreciation for native XML databases at a recent trade show, for which his company decided to alter an application that collected patient and medical records so that the underlying database was Tamino rather than a relational database. The applications consisted of a back-end, where streams of patient data and medical records were arriving in non-XML format, being converted, and then stored to the database. A front-end Web client, meanwhile, was called upon to search and query documents, and return the information as formatted HTML.

"Here we were, storing BLOBs and index fields, and in general, doing a whole lot of things that wanted to be XML, but weren't," recalls Brown. But by moving to Tamino, "we cut our work way down." Performance was better, for one, but more to the point, "it beat pulling the BLOB, interacting with it, serializing it, and putting it back in. It was just a much cleaner process."

FiveSight is actively evaluating a number of different XML database products, in response to what Brown describes as customers who are increasingly aware of the XML database issue. "Customers regularly come to us now and ask 'What sorts of XML databases do you support?'" Brown says.

Native How?

But what exactly does "native" mean in the context of an XML database? In practice, it might be taken to mean any database that stores XML but that does not have an underlying relational structure. Things are not that simple however. B-Bop, for example, offers its Xfinity XML database as a native XML database that relies on a relational database for the underlying storage, specifically, Oracle or Microsoft SQL Server.

How does B-Bop justify this decision? According to Sengupta, even so-called "native" XML database vendors engage in some sort of mapping. "The only really native way of storing XML is as a string in a file system," he points out.

In his article on XML Databases (`http://www.rpbourret.com/xml/XMLDatabaseProds.htm`), XML developer Ron Bourret describes the differences between native XML and XML-enabled (e.g., relational) databases as threefold. To paraphrase: Native XML databases can preserve physical structure (e.g., entity usage, CDATA sections); native XML databases can store XML documents without knowing their schema (DTD), or even schema-less documents; and interaction with a native XML database occurs via XML

Server/Workstation
EXPERT
UNIX
Citrix MetaFrame for UNIX
Server/Workstation
EXPERT
Demystifying OBJECT Options
Server/Workstation
EXPERT
INTEGRATING ERP and CRM VIA THE WEB
Server/Workstation REPRINTS
Keep up with the latest in server and workstation technology with the best minds in the industry. Use reprints to promote, inform, and sell.
Reprint Management Services provides you with the opportunity to obtain reprints of reviews, articles, and features in Server/Workstation Expert.
High-quality editorial reprints will help your company in many ways:
Increased exposure for your product or service
Credible, believable information that consumers trust
Excellent sales tool for mailings, trade shows, and media kits
Powerful educational resource for consumers and employees
Reprints can be completely customized to fit your companies' needs.
Call today for additional information!
Cathy Hensel
Reprint Operations Specialist
Phone (717) 399-1900
Fax (717) 399-8900
REPRINT MANAGEMENT SERVICES
1811 Olde Homestead Lane
Lancaster, PA 17601
Email: sales@rmsreprints.com
Server/Workstation
EXPERT
2001: A Developer's Odyssey
Server/Workstation
EXPERT
Server/Workstation
EXPERT

and XML-related technologies, for example, the Document Object Model (DOM), Xpath, or an XML-based API, such as the XML:DB API.

From B-Bop's perspective, then, the important thing to consider when choosing an XML database is functionality. "What we mean by native is that is has all the capabilities that an XML database should provide," says Sengupta. "What is used underneath is really immaterial."

XYZFind software developer Evan Lenz agrees with this assessment: "By native, I mean not how the database is implemented under the covers, but the interface you give to the user."

XML and Databases

Obviously, what to look for from an XML data storage solution depends entirely on the application being developed, and the organization–its existing infrastructure, its business partners, its personnel, its budget, etc. But, in what kind of applications are developers using databases to store XML?

As a general rule, it's pretty safe to say there are two basic scenarios where XML is being used, says Grabscheid.

The first, and still most prevalent, says Grabscheid, is to use XML as a data exchange mechanism.

InterSystems customer Partners Healthcare, in Boston, MA, for example, uses XML to create an illusion of a single coherent system between all of its disparate platforms, databases, and applications, says Steve Flammini, director of application development. For example, a doctor might go up to a workstation looking for a patient's latest lab results. Those results might link to an image processed by a different application from the radiology labs. "The findings and images are integrated together, even though they come from disparate applications."

To do this, Partners relies on the Cache post-relational database and application server. Data is not stored natively as XML, but rather parsed in and out of Cache's object model using an XML DOM parser from Microsoft Corp., as well as one developed in-house. "There's really no benefit for us to store the documents as XML directly," says Flammini, "it would actually require more overhead for us to do so."

The second application for XML–and this is largely where native XML databases come into play–is to apply XML to various document management functions–for example, Web publishing and content management. Developers of these systems will relish the easy search, query, and transformation enabled by storing data as XML directly within a database.

The State of Georgia recently unveiled the New Georgia Encyclopedia, for example, an online collection of XML-based documents relating to Georgian culture and history that uses XML Information Server (XIS) from eXcelon Corp., formerly ObjectDesign. The encyclopedia was designed by Atlanta, GA-based Merrill Hall, which created the front-end–"a typical Web encyclopedia," according to solutions architect Ben Hall–as well as a Java-based administrative tool allowing the encyclopedia's editors to categorize, aggregate, and capture metadata on the encyclopedia content.

Merrill Hall looked into developing the encyclopedia around Oracle 8i, Hall says, but ultimately decided to go with XIS for a few compelling reasons. Initial prototyping of the encyclopedia using Oracle 8i revealed some "serious performance issues," says Hall, and mapping the document's highly irregular structure in a relational format was "awkward at best." [Oracle 8i] "had promised a lot, but it didn't deliver," says Hall.

On the other hand, storing the XML data in XIS "gave the users a lot of different ways of getting at the data," says Hall, without the headaches and performance issues of mapping and parsing to and from a relational schema.

The question of what kind of database to use can be phrased another way, according to Bourret, whose Web page (`http://www.rpbourret.com`) features an extensive list of XML database products and a discussion of issues facing XML database storage: "Is my XML data- or document-centric?" is the pivotal test developers should put to their XML. That is, is your XML highly structured, with unchanging, regular fields, in the manner of a purchase order. Or, is your XML "semi-structured." That is, is it a combination of structured data with freer-flowing content alongside it–an encyclopedia entry, for example, accompanied by images and audio clips?

Where your XML falls within that spectrum will largely determine what you are looking for from an XML database storage solution.

Where your XML falls within that spectrum will largely determine what you are looking for from an XML database storage solution. Developers using XML to exchange data with business partners or between applications, for example, will probably content themselves with XML utilities provided in traditional transactional database products.

Things to Look For

So what are the specific things that developers should be looking for from a XML database storage solution?

From a developer's standpoint–especially one working with rich XML data–perhaps the biggest concern is how easy it is to interact with the database.

Because while it may be fine to map data once or twice, remapping data further on down the road can get pretty tedious, developers say. "The time people are really going to love their XML databases isn't when they first install them," says FiveSight's Brown. "When they'll really pat themselves on the back is when the time comes to go back and

change the way external applications interact with the data."

Take for example, XYZFind: Based on search technology called inverted indexing, its central design requirement, says Lenz, "is to make the interface as simple as possible for the user." To that end, it takes the approach of indexing every aspect of the XML document without requiring a schema or a DTD. "There's no configuration required. It's as simple as storing the data on the file system, but you can still do structured query and search," says Lenz.

For many developers, ease-of-use is gauged by how extensively a database supports emerging XML standards. "Standards compliance is the one thing that lets me defeat *caveat emptor*," says FiveSight's Brown. "If a vendor says the product supports DOM Level 1, I know exactly what I can do with it."

Among the standards initiative relative to XML and databases are programming interfaces: DTDs, the DOM (Document Object Model), SAX (Simple API for XML, or XML Schema); query languages (XQuery, XPath, XQL); and transformation tools such as XSL (eXtensible Style Language) and XSLT, (eXtensible Style Transformation Language). In the case of relational databases, it's also relevant to ask if the product supports ANSI's recent XML extensions to SQL.

Especially in the area of query languages, XML standards support is a work in progress. Under the auspices of the W3C's XML Query Language Working Group, developers are hammering out the details of several standards, but "it still has a way to go before it achieves a single standard query language," says Philippe Vauclair, vice president of business development for IxiaSoft. One of the earlier query languages, the SQL-like XQL, is losing in popularity, and developers are turning toward XPath, a language that lets you address subsets of a document; and XQuery, which allows you to address collections of documents.

But like with all standards, it's important to pin down exactly what a vendor means when it says it supports this or that standard. "Most of our competitors say they support XPath," says XYZFind's Lenz, "so our customers come to us and say 'Why don't you support XPath?' We always say to them, 'Take a closer look.' It turns out that XPath support is always a qualified thing."

A Moving Target

But just as much as XML standards are in a state of flux, so is the database industry's support for XML. Increasingly, it's becoming harder to say who's native and who's not.

This June, Oracle finally started shipping its much ballyhooed Oracle 9i database, which, among other things, features enhanced XML capabilities.

Not least among the XML-related enhancements is the addition of a new object XML data type, XMLType, which effectively promotes Oracle 9i to the status of native XML database.

Sybase ASE 12.5, released in June, also claims native XML support via an XML data type. And while IBM has not yet added an XML data type to DB2, "I suspect we might have to," says IBM's Selinger. "It's a matter of efficiency."

Also new in Oracle 9i is a pre-loaded C XML Developer's Kit (XDK) that allows developers direct access to XML-related technologies such as XML/XSLT parsers, XML Schema parsers, XML Class Generators, and an XSQL Servlet. The XDK is also available in Java, PL/SQL, and C++ versions.

So, as relational database vendors get increasingly sophisticated with their XML support, Oracle's Magee's sentiment that "customers won't move their data and infrastructure unless they absolutely have to" resounds more loudly.

Still, the Yankee Group's Rob Perry feels that start-ups focusing on native XML database storage stand a good chance of success, "especially in the realm of document management."

Because, while relational database vendors may inch closer and closer to native status, the fact of the matter is that a native XML data type "is still an option that requires a lot of work," says XYZFind's Lenz. "That wouldn't be good enough for me. If there's a product out there that requires no work, and delivers all the functionality you need, doesn't that seem more advantageous?"

Companies Mentioned in This Article

B-Bop Associates Inc.
1350 Bayshore Highway
Suite 100
Burlingame, CA 94010
http://www.b-bop.com

eXcelon
25 Mall Road
Burlington, MA 01803
http://www.exceloncorp.com

FiveSight Technologies Inc.
213 N. Morgan, Suite 1A
Chicago, IL 60607
http://www.fivesight.com

IBM Corp.
Contact local sales office
http://www.ibm.com

InterSystems Corp.
One Memorial Drive
Cambridge, MA 02142
http://www.intersystems.com

Ipedo Inc.
1001 Marshall St.
Redwood City, CA 94063
http://www.ipedo.com

Ixiasoft Inc.
825 Querbes Ave., Suite 200
Outremont, Quebec H2V 3X1
Canada
http://www.ixiasoft.com

Microsoft Corp.
One Microsoft Way
Redmond, WA 98052
http://www.microsoft.com

Oracle Corp.
500 Oracle Parkway
Redwood Shores, CA 94065
http://www.oracle.com

Software AG
Uhlandstr. 12
64297 Darmstadt
Germany
http://www.softwareag.com

Sybase Inc.
6475 Christie Ave.
Emeryville, CA 94608
http://www.sybase.com

XYZFind Corp.
1607 13th Ave.
Seattle, WA 98122
http://www.xyzfind.com

Traffic Analysis on the Threes

by **IAN WESTMACOTT, Technical Editor**

Three vendors vie to be your "eye in the sky" for Web site usage.

This month I look at three log analysis and reporting packages: Quantified Systems' Urchin, Sane Solutions' NetTracker, and WebTrends' Enterprise Reporting Server. All three products offer multisite log file analysis with drill down detail, Web-based interfaces, and multiplatform support.

To test these products, I set up a multisite environment consisting of three Web sites and a search engine. I collected about 335 MB of log data (in extended log file format) for the four hosts containing a total of 1.5 million entries. Because it is common for Web servers to sometimes write invalid log file entries, I threw some of those in the test files as well.

For each product I configured a test run that created individual reports for each of the four sites and a summary report for the combined sites. The log processing host was a Sun Enterprise 250 with two 400 MHz processors, 1 GB real memory, 2.5 GB swap, and 10,000 RPM drives. The same host provided DNS (Domain Name Services or Systems) resolution, using the 9.1.2 release of BIND. The processing time of these analyzers is dominated by DNS resolution. The performance of your name server is at least as important, if not more so, as the performance of the log processing host itself (see "Notes on Benchmarking," Page 54).

In testing the reports produced by the products, in addition to looking at the standard reports, I asked if I could produce two custom reports. In the first, I wanted to know how many users referred by host X eventually bought product Y. For this report, the analyzer must be able to track paths through the site which have a particular starting point outside the domain and end, or include a particular page. In the second report I wanted to find the most popular paths *across* all four test sites. That is, instead of just tracking user paths within each site, I wanted to track paths which moved between sites. For this report the analyzer must be able to analyze multiple sites as a group.

Quantified Systems' Urchin 3 Dedicated

Quantified Systems Inc. was formed in 1995 as a Web hosting and development company. Urchin was developed as an internal tool to help manage Quantified's sites, until the company decided to productize it. Urchin 3 is available in four editions: Pro (single site, $199); Dedicated (multisite, $495); Multihome (clustered, $5,000); and Enterprise (co-brandable, APIs, engineering support, contract pricing). Urchin Dedicated is also available for Sun's Cobalt servers, integrated with the Cobalt management interface. This review covers Urchin 3 Dedicated, which is available for Windows 2000/NT 4, AIX, BSDI, Cobalt, FreeBSD, HP-UX, IRIX, Mac OS X, Red Hat Linux, Solaris, SuSE Linux, Turbolinux and Yellow Dog Linux.

Features

Urchin 3 log processing is command-line driven (except on the Cobalt) and is guided by a text-based configuration file (there is no administrative HTML interface). Urchin can be configured to perform log rotation (including restart of the Web server) or deletion and can be scheduled with `cron`. Urchin does not serve reports itself–you must have and configure a Web server to view the reports. Because Urchin produces dynamic reports (processes them at viewing time), this Web server must support CGI and symbolic links.

Urchin gives you control over how it performs DNS resolution: you can turn it on and off, you can choose the DNS server to use (it doesn't have to be the same as the server), and you can set time-out thresholds: You can set how many times Urchin loops to resolve addresses, and the minimum percentage of addresses to resolve before giving up.

Urchin provides POSIX extended regular expression filtering, so you can filter out or in certain log file entries (such as exclude local hosts) and capture CGI parameters. For example, if you have a CGI program that has a number of parameters, different values of those parameters will look like different pages which get reported on, even though this is really the same page. Using URL filtering, you can collapse differing parameter values into a single page for reporting purposes. You can also match certain parameters, up to four, to report on separately. For example, customers buying widgets with the form `buywidget.cgi` lets you view how many bought blue widgets (`buywidget.cgi ... ?color=blue ...`) versus those that bought red widgets (`buywidget.cgi ... ?color=red ...`).

An e-Urchin ROIR (return on investment reporting) module is available. It enables reports to be viewed in terms of dollars spent and adds an e-commerce report that analyzes products and regions. The e-Urchin log format (ELF) is supported by Miva's Merchant product.

Usage

Installation of Urchin Dedicated was seemingly straightforward. The software is provided as a compressed tar file; simply un-tar on the system to do the processing (it needs filesystem access to the logs to be processed), edit a configuration file, and configure a Web server (not included) to serve the reports. Documentation is provided as a PDF file, although this contains no information on the produced reports–the Urchin Web site has report documentation.

However, when I tried to access the reports they were inaccessible. I had configured Netscape Enterprise to serve the reports (while the documentation has specific instructions for Netscape and Apache servers, any server supporting CGI and symbolic links should do), but a misconfiguration prevented the reports from being processed correctly. The somewhat terse documentation provided no help for this situation, so I fired off a message to technical support.

That first message to technical support was lost, apparently due to a new tracking system Quantified was working with. I got through four days later on a second try, only to be told that support did not yet have a Netscape test setup and they did not know what the problem could be. I eventually discovered the misconfiguration myself.

Reports

Urchin reports are in HTML format and can be viewed in 10 different languages. A browser that supports JavaScript is required, including Netscape Navigator 3.01 or later and Microsoft Internet Explorer 3.0 J2 or later. The reports are processed at viewing time from the flat-file database and can be exported as tab-delimited data.

Reports are divided into a header frame, a navigation frame and a report frame (see Figure 1). Individual site report navigation is divided into seven sections: Domains, Pages, Referrals, Traffic, Browsers, E-Commerce and Tracking. The E-Commerce reports are useful only with the optional e-Urchin module installed (I didn't install it).

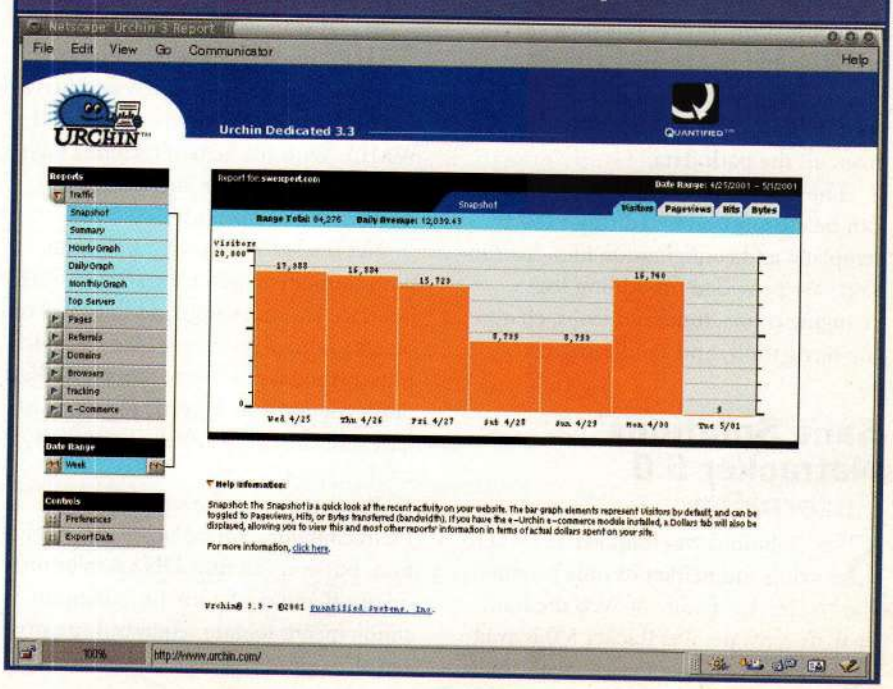
The Traffic section provides a snapshot graph of hourly, daily or monthly traffic for the chosen range in terms of visitors, page views, hits, views or bytes, and a table summary of total and average traffic for the range. The Pages section provides pages, directories and file types statistics in views or bytes as number of views and percentage of total views.

The Referrer section provides referring sites, pages, search engines and keywords in number of visitors, percentage of total visitors, and average number of pages viewed. The Domain section provides domains and top-level domains in number of visitors, percentage of total visitors, and average number of pages viewed.

The Browsers section provides browsers, platforms, and browser/platform combinations in number of visitors, percentage of total visitors and average number of pages viewed. The Tracking section provides entry and exit pages by number of views and percentage of total views, paths taken through the site sorted by entry page in terms of number of views and percentage of total, depth and length of visits by number of visitors, and authenticated users by number of visits and percentage of total.

Urchin is fast, and one of the reasons it is fast is the comparatively limited reports it produces. For example, Urchin will not give you the top hosts visiting the site, just the top domains. Quantified argues that top host statistics are not very informative; we would argue otherwise, particularly in an intranet setting. It would be nice if the report were available, but could be turned on or off.

Figure 1. Urchin Traffic Snapshot Report



Urchin 3 Dedicated

Price:
$495 for 25 sites; $295 for each additional 25 sites

Company:
Quantified Systems Inc.
2165 India St.
San Diego, CA 92101
http://www.quantified.com

Best Feature:
Speed

Worst Feature:
Limited flexibility and reporting

The reports are not customizable beyond date range. For instance, the Tracking report analyzes paths taken through the sites, but is sorted by entry page. Thus you cannot analyze reverse paths, and my first custom report could not be created. Quantified tells us that the e-Urchin module will give you referrers in terms of dollars per visitor, which would indicate the data I was looking for, but is not quite the same.

The system report provides aggregate traffic and bandwidth data for all the individual sites analyzed by Urchin. However, paths cannot be analyzed across sites, so my second custom report could also not be created. You could specify the different site's log files as logs from a single site, but then the reports could not be linked into the sites correctly, and the multisite paths would have to be searched for by hand from all the path data.

The interface of the Urchin reports can be customized by editing HTML templates and configuration files. Instructions are provided for adding logos, changing colors, fonts and icons, changing navigation, and modifying text.

Sane Solutions' Netracker 5.0 Enterprise

Sane Solutions was founded in 1996 to develop and market its only product, the NetTracker family of Web site traffic analysis software. NetTracker 5.0 is available in three editions: Professional (single site, $495); Enterprise (multisite plus FTP, proxy and firewall analysis, $995); and eBusiness Edition (Enterprise plus database integration, $9,995). I reviewed NetTracker 5.0 Enterprise.

NetTracker is available for Windows (NT/95/98/Me/2000) and UNIX (Solaris, Linux, AIX, HP-UX, Tru64, IRIX) platforms. All NetTracker versions support cluster configurations.

Features

NetTracker Professional and Enterprise import log data into an embedded flat file database from which HTML reports are generated, while the eBusiness edition stores log data in an Oracle or Microsoft SQL Server database. Ebusiness sessionizes the traffic data and creates summary and detail tables, based on a documented schema that may be used to integrate the data with CRM and BI (business intelligence) applications such as solutions from MicroStrategy, Business Objects, E.piphany, Seagate Software (Crystal Reports), and Cognos.

NetTracker also supports e-commerce platforms, such as those from Allaire, ATG, BEA, BroadVision, IBM, Intershop, and Vignette, by having the ability to parse parameters and translate cookies. NetTracker includes server plugins for Apache, Netscape, iPlanet, and Microsoft servers that use cookies to log additional information and can be used to report on server performance and page delivery.

In addition to Web site logs, NetTracker (Enterprise and eBusiness editions) can analyze FTP, proxy and firewall log files, and wireless traffic (WAP and WML). With the help of included CGI programs, it can track and analyze ad banners and campaigns.

NetTracker Enterprise can be run both through an administrative HTML interface and the command line, and can be scheduled with `cron`. Like Urchin, Enterprise does not serve reports itself; it produces dynamic reports, so you must have and configure a Web server with CGI support.

Enterprise uses the server's default DNS configuration for hostname resolution, but you can turn DNS resolution on or off and configure the maximum simultaneous lookup requests. Enterprise can also optionally resolve page titles, so that meaningful titles can be used in the reports rather than cryptic URLs.

Also like Urchin, NetTracker lets you use regular expressions to filter data included in custom reports. The NetTracker custom report wizard includes a set of pre-defined filter groups and a simple expression builder that simplifies things for those unfamiliar with regular expressions. In addition, a large number of common filter combinations are available when defining reports, such as grouping, including and excluding classes of entities, analyzing ad campaigns and defining visitor profiles.

You can add arbitrary URL search and replace definitions to pre-process the logs. This feature can be useful for collapsing different URLs into one for analysis purposes, such as URLs that include unique identifiers but are really the same page, or counting all the pages in a directory as one. Rather than rewrite URLs before analysis to group pages (which will affect all statistics), you can define content groups by using regular expressions to match pages and parameters, which will leave the log files as-is and report statistics on each page and on the content groups.

A Department definition and report feature lets you categorize ranges of intranet addresses into organizational units for analysis.

NetTracker can access log files through the filesystem or via FTP and can handle logs compressed with *gzip*.

Usage

UNIX installation is by way of a shell script, which had little error handling. A permissions problem caused the script to fail rather ungracefully. Once that was fixed I was all set, merely specifying program, HTML and CGI directories. A well-written and comprehensive 300-page User's Guide is included. In fact, due to the manual and the smooth operation of NetTracker, I never had a reason to bother technical support. So I made up a question.

I sent e-mail to technical support asking how to create the second custom report. After several hours I received e-mail back asking to clarify the question. Having done that, I received a polite and thorough description by e-mail a day

LinuxWorld™ CONFERENCE & EXPO

Where Open Source Comes of Age

Don't miss the world's leading Linux event focusing on Linux and Open Source solutions. Whether you're looking for a better way to manage and scale your network, or a more cost-effective way to run your company, LinuxWorld Conference & Expo can show you how. Be there to network, learn and see where Open Source comes of age.

- Hear **Keynotes** from leading Linux and Open Source visionaries
  - **Shane Robison**, Sr. Vice President & CTO, Compaq Computer Corporation
  - **Larry Lessig**, Law Professor, Stanford Law School
  - **Matthew Szulik**, President & CEO, Red Hat
- Attend Feature Presentations such as **"Golden Penguin Bowl"** hosted by Nicholas Petreley, **"The State of Open Source"** featuring Linus Torvalds, hosted by Larry Augustin, CEO, VA Linux Systems, and **"Best Practices with Linux for e-Business Infrastructure"** moderated by Ross Mauri, Vice President, IBM and President, OSDN Lab.
- **Experience** an interactive exhibit floor that will showcase the hottest Linux products, services, applications and embedded systems from over 250 leading hardware and software vendors.
- **NEW! The Worldwide Developer Conference at LinuxWorld** geared exclusively to application developers, along with the most comprehensive and very best Linux educational training available in the world.

Conferences
August 26-30, 2001

Expo
August 28-30, 2001

Moscone Convention Center
San Francisco, CA

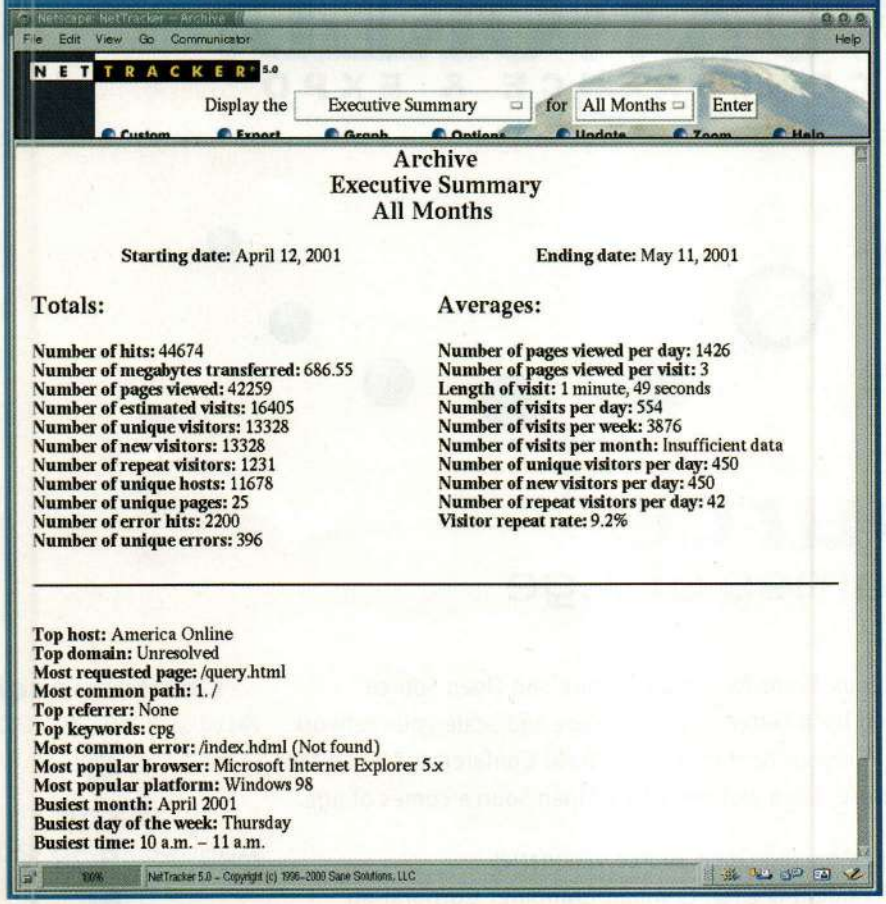
Owned & Managed By:
IDG WORLD EXPO

© 2001 IDG World Expo. All rights reserved. LinuxWorld is a trademark of International Data Group, Inc. All other marks contained herein are the property of their respective owners.

Cornerstone Sponsor	Platinum Sponsors	Certification Sponsor
VA LINUX SYSTEMS	Agenda COMPUTING, COMPAQ, ca Computer Associates, GreatBridge, hp invent, IBM, intel, redhat, SERVEL LINUX, sgi, Sun microsystems, SuSE	LINUX GNU

www.linuxworldexpo.com

Figure 2. The NetTracker Executive Summary Report


File Edit View Go Communicator Help
NETTRACKER 5.0
Display the Executive Summary for All Months Enter
Archive
Executive Summary
All Months
Starting date: April 12, 2001
Ending date: May 11, 2001
Totals:
Number of hits: 44674
Number of megabytes transferred: 686.55
Number of pages viewed: 42259
Number of estimated visits: 16405
Number of unique visitors: 13328
Number of new visitors: 13328
Number of repeat visitors: 1231
Number of unique hosts: 11678
Number of unique pages: 25
Number of error hits: 2200
Number of unique errors: 396
Averages:
Number of pages viewed per day: 1426
Number of pages viewed per visit: 3
Length of visit: 1 minute, 49 seconds
Number of visits per day: 554
Number of visits per week: 3876
Number of visits per month: Insufficient data
Number of unique visitors per day: 450
Number of new visitors per day: 450
Number of repeat visitors per day: 42
Visitor repeat rate: 9.2%
Top host: America Online
Top domain: Unresolved
Most requested page: /query.html
Most common path: 1. /
Top referrer: None
Top keywords: cpg
Most common error: /index.hdml (Not found)
Most popular browser: Microsoft Internet Explorer 5.x
Most popular platform: Windows 98
Busiest month: April 2001
Busiest day of the week: Thursday
Busiest time: 10 a.m. – 11 a.m.
NetTracker 5.0 – Copyright (c) 1996–2000 Sane Solutions, LLC

later, along with a phone number to follow up if necessary.

The HTML administration interface lets you perform all configuration and analysis through a Web browser. This is convenient for remote administration functionality but, like most HTML/CGI administration interfaces, application errors are not uncommon. Thankfully command-line use is well documented.

Reports

NetTracker Enterprise can create both individual and overall reports, and custom reports can be defined. There are 86 individual reports (35 Web, 18 proxy and 13 FTP) and 20 overall reports (nine Web, six proxy and five FTP). Reports are generated at viewing time, and can be viewed in English, German, French, Spanish and Japanese. Netscape 3.x and 4.x and Microsoft IE 3.x, 4.x and 5.x are supported, and you can configure Enterprise to e-mail reports, as well as export data in HTML, comma-separated format, or RTF. Graphs are created with a Java applet that lets you graph data in various 3D graphs and zoom in or out.

NetTracker reports are divided into a navigation frame and a report frame. The navigation frame lets you choose the report and a monthly time period (Figure 2). The report is then displayed in the report frame. The report frame includes a regular expression builder that lets you filter the items in the report.

It became clear that Sane Solutions' strong point is not HTML design, as the reports were rather spartan and with layout difficulties. For example, using Netscape Navigator 4.7 on my SPARC/Solaris workstation resulted in misaligned frames and background images.

NetTracker supports an interesting feature that lets you specify lookup URLs for cookies, hosts, users and visitors. These URLs specify CGI programs which can query databases to extract extra information. If you specify such URLs, NetTracker will link them into the reports, so you can drill down into your database.

The 35 standard reports include temporal summaries showing views and visits by date, day of week and time of day; visitor summaries giving views, visits, time online, and last visits by visitors, number of visits by duration time, unique, new and repeat visitors, number of visitors by number of visits, visitor profiles when using cookies, and authenticated users; client summaries reporting platforms and browsers, hosts and domains, and a Robot/Spider summary giving statistics for common spiders; click-path summaries including entry and exit pages, referrers and initial referrers, keywords used to find the site, and links used to leave (using a CGI program); page summaries in terms of views and visits, delivery, local keywords and parameters, directories, by content and by department; traffic, error and server performance summaries; cookie and ad campaign summaries; and an executive summary giving high-level site usage details.

Overall summaries provide bandwidth, browser, day of the week, hit, platform, time, view, visit and visitor summaries for a group of sites.

Custom reports are created with a wizard that lets you start from a standard report and customize it by applying filters and defining the number of entries to show per page. There is no documentation for modifying the design of the reports or rebranding them.

My two custom reports were simple with NetTracker. In fact, the first one is given as an example of how to create custom reports in the documentation.

NetTracker 5.0 Enterprise

Price:
$995

Company:
Sane Solutions LLC
35 Belver Ave., Suite 230
North Kingstown, RI 02852
http://www.sane.com

Best Feature:
Flexible reporting options

Worst Feature:
Interface design

You create a custom report based on the Referrer Summary report, adding a page filter to include only pages that match the product Y purchase page. Then you run the report and filter the results on referrer X.

To achieve the second custom report, I first had to define a new site to analyze. NetTracker can analyze sites based on a single log file (Not Clustered), multiple log files for a single site (Clustered for a Single Site), or multiple log files for multiple sites (Clustered for Multiple Sites). Using the latter, NetTracker is able to analyze multiple sites simultaneously, and all 35 standard reports are created. All URLs in the reports are correctly prefaced by the sites from which they came, and NetTracker was the only analyzer that could correctly analyze multiple independent sites as a group.

However, the Paths report for the multiple-site analysis includes all paths, including ones within a single site, so you must tease out the multisite paths by hand.

WebTrends' Enterprise Reporting Server 3.5

WebTrends launched its first product, WebTrends Log Analyzer, in 1994. Today, WebTrends offers a large array of products: Log Analyzer (log analysis without report serving, $699); Security Analyzer (security assessment, $1,995 for 255 addresses; $4,999 for unlimited addresses); Firewall Suite (firewall/VPN/proxy server analysis, traffic categorizing, no serving $2,499); Professional Suite (log/streaming media/proxy/link analysis and alert/monitoring of 10 devices without report serving, $999); Enterprise Suite (Professional Suite for unlimited device monitoring $2,499 plus $1,799 per site); and Enterprise Reporting Server (includes report serving, $5,998 Solaris plus $3,598 per additional server; $4,199 Linux/NT plus $2,399 per additional server). The ERS and Suite products include ClusterTrends, technology that combines multiple logs for cluster analysis. It is licensed separately. The CommerceTrends Family, or "VRM platform," provides data integration and analysis for ERP/CRM applications, such as Vingette, Allaire, BroadVision, and OLAP ($100,000 and up). WebTrends Live is a hosted version of ERS with some additional functionality from some of the other products.

I reviewed Enterprise Reporting Server. It is available for Windows NT 4.X, Red Hat Linux 5.2, and Solaris 2.6, 2.7 and 2.8. ERS is licensed for 500 domains and five end-users and a single physical machine.

Features

ERS includes several attractive flexibility features. It lets you choose whether to pregenerate reports or create on-demand reports, to specify arbitrary pre- and post-processing applications to use, and to control resource usage and performance in several ways. For example, you can control memory usage by limiting the size of certain statistics tables (such as how many top documents to track), you can set the maximum number of concurrent analyses to perform, and the maximum number of days to process at a time.

ERS also provides a "character analysis" that you can run on profiles, which produces a detailed report of the log analysis, including resource usage, errors and counts. This can be very useful for debugging log files and profiles.

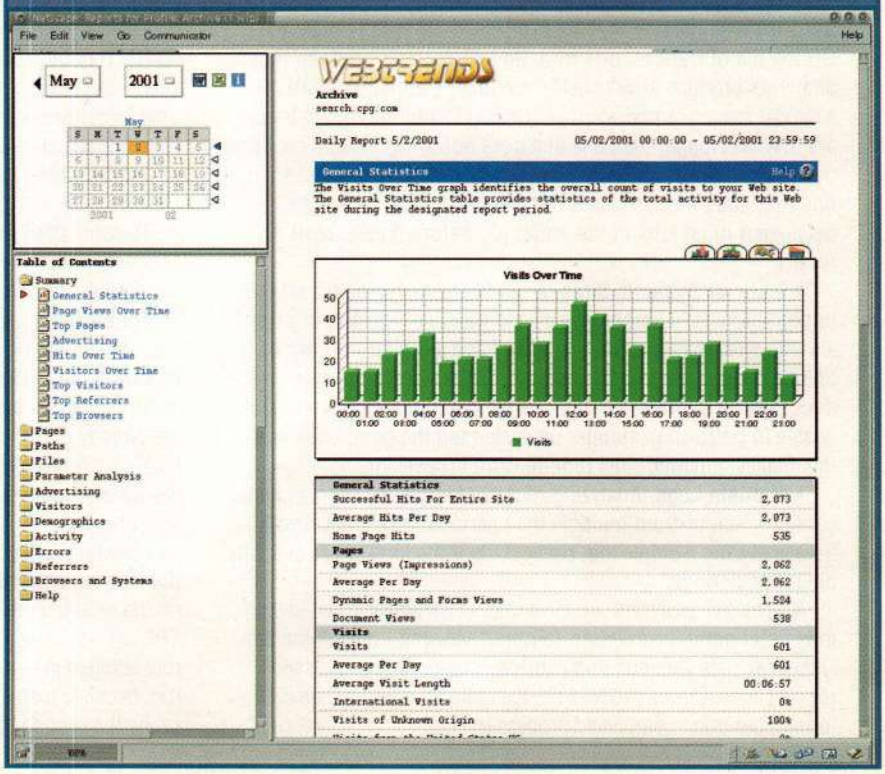
DNS lookup can be enabled or disabled, and you can set the number of lookup threads to use and the size of the lookup queue. Page titles can be optionally resolved, and ERS caches both page titles and DNS lookups–a nice feature.

ERS can perform URL and cookie filtering and text search to perform analysis on subsets of the logs. WebTrends provides server plugins for Netscape and Microsoft Web servers that implement cookies for user tracking analysis (download and documentation at the WebTrends Web site), or you can use your own. ERS has a wizard for defining profiles that track ad views and clicks (essentially by just filtering the logs on the ad filename and clickthrough text), and reports for the analysis.

As with NetTracker, you can apply arbitrary URL search and replace definitions and define content groups. Also as with NetTracker, you can define departments or organizational units by address.

In addition to tracking top paths through the site, ERS allows you to define particular paths to analyze. This is accomplished by choosing a particular

Figure 3. ERS General Statistics Summary Report



page to track and the number of levels to and from the page to track. However, you cannot specify other characteristics of the tracked page, such as whether it is an entry or exit page for the visit.

ERS can also track parameters in URLs, such as form data sent to CGI programs. Each parameter analysis definition can track up to two parameters. You can define translations of cryptic parameter values to meaningful phrases for reports.

ERS can retrieve log files through the filesystem or FTP. ERS will not do log archiving or rotation itself, but you could use the pre- and post-processing features to configure this yourself.

Usage

On Solaris, ERS can be installed either with a script or as a Solaris package. ERS can be configured and controlled through the Web with its included HTTP server, or on the command line (and configuration files are plain text). In normal operations, ERS runs as a daemon listening for requests on a network port and processes log files according to a scheduled process queue. The administrative Web server communicates with ERS via the network port, sending commands and polling for status.

As with all HTML/CGI interfaces, applications errors are not uncommon. For example, I found that sometimes setting options through the administrative server did not take effect. But the server can also run in non-daemon mode for scripted applications, which is well documented. Still, of the three products, ERS had the best HTML interface.

Documentation is by way of a 300-page booklet. The booklet covers most of the areas you would want, including command-line operation and basic troubleshooting, but it is fraught with errors. Pathnames and commands are wrong (which can be very troublesome for novices), page references are wrong, and the document writers have all sorts of notes to themselves which they forgot to remove before printing ("*josh, I have conflicting sets of info: which is correct?*"). The index, which has alphabetical headings, is not alphabetical and has formatting problems.

I got an interesting result for log file with errors. This log file contained a large section of bad lines, lines which could not be recognized as log file entries. These lines were in the middle of the file, with good entries before and after. When ERS encountered this bad section, it just stopped. It did not skip all the bad ones to find the good ones at the end, nor did it give any indication of the problem. It just produced reports which covered the data at the beginning of the file before the bad section.

Notes on Benchmarking

Benchmarking log analyzers is not straightforward. Simple processing speed (in terms of MB/s), and some vendors report this, is a meaningless metric to evaluate performance for several reasons. First, no two log files are alike. There are a number of data factors which may affect the performance of a log analyzer:

- **Temporal Range and Order**. A 1 GB log file of transactions for a single day may be processed differently than a 1 GB log file of transactions from an entire week, because most analyzers produce daily (weekly, monthly) summary reports. For a similar reason, a two-week period that includes a month boundary may take longer than one that does not. A log file that contains out-of-temporal-order entries will take longer to analyze than one that does not, provided the analyzer can even handle it, because it must look at the entire log before it can create any report.
- **Static or Dynamic Sites.** A log file from a dynamic site will undoubtedly take longer to process than a same-size log from a static site. Dynamic sites usually have a higher number of unique pages (due to parameters in URLs), and analyzers produce statistics for each unique page. In addition, if the analyzer is able to perform parameter analysis (and the good ones are), this makes dynamic sites require extra processing.
- **Multiple Logs.** Analyzing multiple logs, such as in a cluster configuration or from multiple time periods, takes more work than analyzing a single log, particularly if the logs cover overlapping time periods.

No two log analyzers are the same. In particular, an analyzer may defer some processing to report viewing time. Some analyzers process the logs and produce static, fully processed report pages. Others process the logs into an intermediate data format, but don't completely process the individual reports until someone actually tries to view them. This makes report viewing slower, but speeds up log processing time. You need to account for this when comparing log analyzers. Moreover, different analyzers may produce different sets of reports, which may affect analysis time.

Log analysis performance is very much dependent on DNS performance. Any good log analyzer provides reverse DNS lookup functionality–the ability to resolve an IP address to a hostname. All HTTP traffic contains IP addresses, not hostnames. So a Web server will store IP addresses in its log files, unless it is configured to perform reverse DNS lookup. Reverse DNS lookup can slow a Web server down, so a better solution is to log IP addresses and then have the log analyzer perform reverse DNS lookup at analysis time.

Reverse DNS lookup can be very slow, so slow that it can account for 50% to 90% of log processing time. Log analysis vendors make claims about optimizing DNS lookup, but at the end of the day they all have to contact a DNS server, and are thus dependent on the performance of that server. The performance of the DNS server is affected by the server software itself, of course, but also its cache (how many of the addresses are already in the cache) and its locality (is the server on the local or a remote machine, and if remote, what is the speed of the network connection to that machine and its processor(s)?).

For example, in my tests I found that using a remote DNS server with a fast connection (100 Mb/s) but slower processors than the log analysis machine was faster than using a DNS server on the local log analysis machine. I attribute this to contention for CPU cycles. I also found that analyzing the same log twice in a row resulted in a 50% increase in performance for the second run, because many of the looked-up addresses were in the DNS cache the second time around.

Table 1. Log analyzer performance

	DNS lookup	No DNS	Size of data/reports
Urchin Dedicated	3:33.7	0:08.7	3 MB
NetTracker Enterprise	4:49.9	3:39.8	1.5 MB
WebTrends ERS	6:11.4	5:53.6	0.75 MB

I ran each log analyzer on a 10 MB log file containing 45,000 entries covering 30 days from two months, both with and without DNS lookup. The results with DNS lookup were from a second run after the DNS results were cached. The data/reports size results were with all analyzers configured to create on-demand reporting. These results should not be taken as an indicator of general performance, but are for comparison purposes only on our particular test setup.

I used this opportunity to contact technical support and was pleased at the immediate delivery of an automated e-mail containing a trouble ticket number. This was followed up several hours later by a courteous and professional support engineer. Unfortunately, the logfile sent at his request was apparently lost and I didn't hear from support for four days when I sent another inquiry. Despite this glitch, WebTrends technical support was first-rate.

Reports

WebTrends ERS reports can be viewed with Netscape Navigator 4.5 or later (except version 6) and Microsoft Internet Explorer 4.01 or later. Your browser will need to support frames and Java. In addition, each report includes icons to view it as a Microsoft Word or Microsoft Excel document, provided you are on a Windows platform (makes use of ActiveX).

The reports are divided into three frames, one to choose the date range and period (daily, weekly, monthly, quarterly and yearly), one to choose the report to view for that date range, and the last to contain the report itself. Each report includes a "help card" containing help information and definitions of terms in the report–a nice feature.

Like NetTracker, all graphs are rendered with a Java graphing applet, which gives you the option of viewing the graphs in various types. When report profiles are created you can add a title, description and graphic that will be included at the top of the report.

Default ERS reports are grouped into Pages, Paths, Files, Visitors, Demographics, Activity, Errors, Referrers, and Browsers and Systems. Most statistics are given by both visits and views, and can be graphed over time. The Pages group gives you top pages, entry and exit pages, least requested and single access pages, top documents and forms (see Figure 3, Page 53).

The Paths group contains a single report ranking the most popular paths through the site. The Files group gives the top directories, accessed and uploaded file types, as well as the top entry and least requested files.

The Visitors group reports on the top visitors and authenticated visitors, and new versus return visitors (provided you are using cookies to track visitors). The Demographics group breaks out statistics by geographic region and IP domain.

The Activity group reports on most or least active dates and days of the week, and daily average hits and visitors. The Errors group reports on errors, while the Referrers group reports on referring sites, URLs and search engines (including keywords used). Finally, the Browsers and Systems group gives statistics on client platforms.

For customized reports, ERS provides profile "templates." These templates let you choose which of the standard reports to include, whether to include help cards and graphs, default graph types, and various default settings such as number of table rows for particular statistics. You can also choose one of a handful of CSS stylesheets for the report, which essentially just change the colors (presumably you could define and add your own stylesheets, but there is no documentation for this).

You can also choose the language, but the only option is English–I have the first release of this version and subsequent releases will include French, German and Spanish. Other languages (including Asian languages) will be available as an add-on.

To create the first of the two custom reports–how many visitors referred by site X bought product Y–is fairly easy using a path analysis definition in combination with a filter. I first define a path analysis definition to track the product Y purchase page with some large number of levels back to track, then filter the log for referrals from site X.

The second custom report, top paths across sites, was not so easy. ERS does not have a special profile for analyzing multiple sites (or multiple profiles); you simply configure a profile with multiple log files. However, unless the individual log files include host information (Multihome; and most do not by default), ERS simply lumps the logs together as if from a single site. The difficulty with this is that you cannot get the reports to link to the multiple sites correctly (because you can only specify one site URL for each profile).

However, you then get the default Top Paths report for the multiple sites, which is similar to the information provided by NetTracker for this report–you must tease out the multisite paths by hand. ⊸

Enterprise Reporting Server 3.5a

Price:
$5,998 Solaris plus $3,598 per additional server; $4,199 Linux/NT plus $2,399 per additional server

Company:
WebTrends Corp.
851 SW 6th Ave., Suite 1200
Portland, OR 97204 USA
http://www.webtrends.com

Best Feature:
HTML interface, both administrative and reports

Worst Feature:
Price

NEW PRODUCTS

The product descriptions are compiled from data supplied by the vendors. For more detailed information, log on to their Web site.

Document Management from Wireless Devices

Organizations using document management solutions such as DOCSOpen, PowerDOCS, DOCSFusion and CyberDOCS from Hummingbird Ltd. can now provide end-users with real-time access to documents from wireless devices such as digital mobile phones, personal digital assistants (PDAs), pagers, and pocket PCs. The new functionality comes by way of a partnership with Matrix Logic Corp. and the integration of its wireless document management system, WirelessDMS.

Based on Matrix Logic's WAPDocs, a document management connector for digital mobile phones, WirelessDMS enables wireless users to access document repositories and convert documents into readable text. Users may also search their document libraries, and then e-mail or fax documents from their handheld devices, Matrix Logic explains. Instant notification of changes to projects, quick searches and documents within the DMS is available via e-mail or directly on end-user's handheld devices.

WirelessDMS works on WAP (Wireless Application Protocol)-capable devices, including a wide range of phones, pagers, Palm devices and Pocket PCs. For mobile users with a Pocket PC or Windows CE handheld, full document synchronization is available when the device is docked.

Matrix Logic Corp.
695 DeLong Ave., Suite 101
Novato, CA 94945
http://www.matrix-logic.com

Internet Security Tool Upgrade

Zone Labs Inc. has announced the latest versions of its Internet security tools: the ZoneAlarm 2.6 and ZoneAlarm Pro 2.6.

New features in both products include a tutorial, online help, color-coded alerts, and security at the OS level, which helps safeguard users from security sabotage.

Both products include a firewall; "Application Control," which lets users control which programs can communicate with the Internet; and MailSafe-Email Attachment Protection.

ZoneAlarm Pro comes equipped with advanced configuration utilities for more technically-knowledgeable users, including an enhanced version of MailSafe; password protection; and support for LANs.

Both products are compatible with Windows 95/98/NT/2000/Me and are free for personal and non-profit use.

ZoneAlarm is $19.95 for business use, while ZoneAlarm Pro costs $39.95. Both are available from Zone Labs' Web site.

Zone Labs Inc.
1060 Howard St.
San Francisco, CA 94103
http://www.zonelabs.com

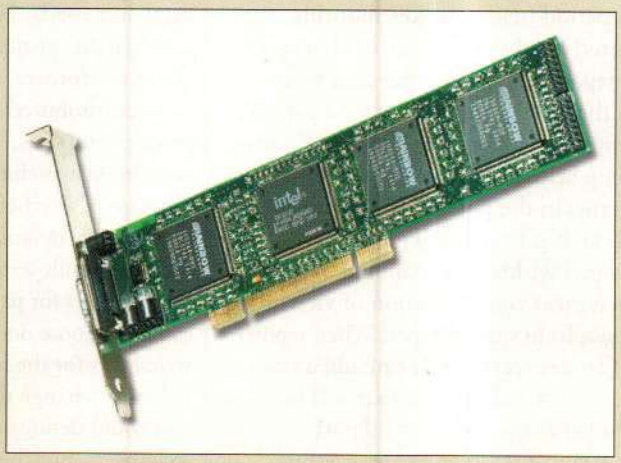
PCI Card Accelerates SSL Sessions

The CryptoSwift 400 eCommerce Accelerator, released by iVEA Technologies, a division of Rainbow Technologies, is a PCI-based card that accelerates Web transactions secured using the Secure Socket Layer (SSL) protocol. With it, iVEA says, e-commerce sites using SSL can handle more traffic securely, while minimizing time users spend waiting.

The CryptoSwift 400, which is benchmarked at 400 SSL transactions-per-second (TPS), relies on a public-key processor that offloads and speeds up the operations which would otherwise saturate the Web server's main processor. The CryptoSwift 400 provides a 1,024-bit RSA transaction processing time of less than 4.9 msecs, according to the company.

The CryptoSwift 400 plugs into the open PCI slot on most Web server hardware platforms, and can be integrated with standard secure Web server software packages including Netscape, iPlanet, Apache, and Microsoft IIS. In addition to SSL, the CryptoSwift 400 also accelerates other popular security protocols such as IPSec, SET, SSH, and TLS, on popular Web server platforms such as Apache, IIS, Netscape, and iPlanet, to name a few. The price for the CryptoSwift 400 eCommerce Accelerator is $3,989.

Rainbow iVEA
8 Hughes
Irvine, CA 92618
http://www.ivea.com



What's New, What's Next, What Works.

Look for It At

Internet World Fall 2001

Today, businesses are looking for practical solutions to help increase productivity, cut costs, become more customer focused and in the end, increase profits. Companies will continue to look for the latest, most relevant technologies.They will look for technologies that offer the best solutions to solve their business problems and also create new opportunities.

Internet World Fall is where your business can find those solutions it's looking for, through its in-depth conference program and exhibit floor brimming with companies showing the latest technologies. Join us at Internet World Fall 2001 and discover what's new, what's next and most importantly, what works — for your business.

THE NINTH ANNUAL

INTERNET WORLD
Fall 2001

Co-Located With:

Storage Area Networks Conference & Expo

ByteandSwitch

LIGHT READING brought to you by Light Reading

eCRM FALL

Freedom Technology Media Group

For more information,
visit www.internetworld.com.
For information about exhibiting,
e-mail: eventsales@iw.com
or call 203-559-2866

Internet World Fall 2001
Conference: October 1-5
Exhibits: October 3-5
Jacob K. Javits
Convention Center
New York City

Produced by:

Penton

Official Sponsors:

INTERNET WORLD

Media Sponsors:

INTERNET WORLD.com

boardwatch

internet.com

PENTON'S m-CommerceWorld.com

Wireless Enterprise • Small Business E-Commerce • Enterprise Streaming • Security • One-to-One Marketing • Extended Supply Chain Services Effective Outsourcing • Content Management • Peer-to-Peer Networking • Advanced E-Business Practices • E-Procurement • Storage Area Networking • Electronic Customer Relationship Management

Code: F01I22

A New Line of LCD Monitors from Sharp

Sharp Systems of America, a Sharp Corp. unit, has announced a new line of six LCD monitors targeted at the corporate workstation market. The monitors incorporate new technology to reduce power consumption and glare and increase resolution and contrast, the company says. All models feature a VESA-compliant mount and dual input capabilities, automatic gain control and Universal Serial Bus (USB) support.

Entry model LL-T1501A ($549) is a 15-inch TFT display supporting XGA (1024x768) resolution. Model LL-T15S1 ($599) includes stereo speakers, and model LL-T1511A ($659) includes Sharp's Anti-Glare Low Reflection (AGLR) technology. AGLR is said to reduce glare 60% over conventional coatings by refracting light at different ratios so that reflected rays cancel each other out.

Model LL-T1610W ($849) is a 16-inch TFT display supporting SXGA (1280x1024) resolution, and includes AGLR and Sharp's Super-V technology. This technology reduces the amount of circuitry space needed, to achieve an aperture ratio of 81%, according to Sharp. Super-V is said to produce contrast ratios of 300:1, increase backlight life to 50,000 hours, and reduce power consumption. The LL-T1610W also features both analog and digital (DVI-D) input, and ICC profile support. The LL-T1810A ($1,399) is an 18.1-inch TFT display supporting SXGA, AGLR, Super-V and ICC profile support. The LL-T2000A ($3,099) is a 19.6-inch TFT display supporting UXGA (1600x1200) resolution, AGLR, Super-V and ICC profile support.

The monitors are being sold through resellers and distributors and will be available through the company's Web site later in the year.

Sharp Electronics Corp.
Sharp Systems of America
Western Regional Headquarters
5901 Bolsa Ave.
Huntington Beach, CA 92647
http://www.sharp-business.com

Lightweight Java Engine

Concentus Technology Corp. has announced the release of the WorkBridge Java Engine, which lets companies define, execute, track and integrate business processes across enterprises and devices.

According to the company, the new tool eliminates fat integration pipes, server databases, and application coding through the support of Web services and the use of XML-based process definitions, which can be executed on any device.

XML-based process documents are made self-executing and users can define them on-the-fly. The document then defines tasks such as who needs to be involved, what resources are necessary and how to route the completed tasks to the next person or application in that process, says the company.

Areas where companies can benefit using the tool include collaborative hubs, P2P processing, PDA-enabled process, and machine monitoring.

Pricing and information is available by contacting the company.

Concentus Technology Corp.
5115 Parkcenter Ave., Suite 150
Dublin, OH 43017-7523
http://www.concentus-tech.com

Content and Code Management Tool

Serena Software has announced the Serena ChangeMan WCM, a tool that manages both Web content such as text and graphics, as well as the code that powers the Web site.

Serena touts ChangeMan WCM as an "end-to-end" solution designed for technical and non-technical users alike. The product automates content-focused processes such as task assignment, resource scheduling, content routing, previewing and editing, and online notification and approvals.

Meanwhile, ChangeMan WCM's software code management facilities allow code to be managed integrally with Web content. The product is integrated with other Serena SCM products, including ChangeMan DS, for Distributed Systems, and ChangeMan ALM, or Application Lifecycle Manager.

The Serena ChangeMan WCM server runs on Windows NT/ 2000. Publishing to Windows, UNIX, Linux, and OS/390 is supported. Pricing is dependent on the number of development seats, and starts at $93,000.

Serena Software Inc.
500 Airport Blvd., 2nd Floor
Burlingame, CA 94010
http://www.serena.com

Windows, UNIX, Linux Management

Heroix has released the Heroix eQ Software Suite, which extends its management reach into Linux environments, and improves cross-platform management with wizard-style tools that guide administrators step-by-step through management tasks, according to the company.

Heroix eQ monitors and manages from tens to thousands of Windows 2000, Windows NT, UNIX, and Linux systems, the company claims, all of which can be viewed from a single Web-based interface. It is designed to monitor a range of cross-platform applications such as Microsoft Exchange, Lotus Notes; databases including SQL Server and Oracle; Web servers such as Apache and Microsoft IIS; as well as other server and network technologies.

Central to Heroix eQ is a so-called purpose-built rule engine that powers the product's system monitoring, problem solving, automated configuration, and self-management capabilities, the company says.

Heroix eQ also implements application auto-discovery, which automatically configures monitoring thresholds appropriate to the normal workload of each application and system. The software automatically detects application and workload changes.

Another Heroix eQ feature–emergency repair–provides GUI-independent commands that let staff reach and repair "hung" Windows servers, or, if the system is beyond repair, provide a clean restart of the system to avoid data loss or corruption, according to the company.

Pricing starts at $695.

Heroix Corp.
120 Wells Ave.
Newton, MA 02459
http://www.heroix.com

Digital Certificate System

RSA Security has announced the latest version of its digital certificate management system, the RSA Keon Certificate Authority (SA) 5.7, which issues, manages, and validates digital certificates.

The new release incorporates RSA Keon OneStep software, which is said to make the deployment of digital certificates transparent to users and easier on administrative staff by letting them take advantage of existing authentication systems.

Enhancements include secure administration, enrollment, directory and logging servers, and a Simple Certificate Enrollment Protocol (SCEP) server that provides automatic enrollment for issuing certificates to SCEP-compliant Virtual Private Network (VPN) devices.

Other features include a signing engine for digitally signing end-user certificates and system events, and an integrated data repository for storing certificates, system data and certificate status information.

Contact company for pricing.

RSA Security Inc.
20 Crosby Drive
Bedford, MA 01730
http://www.rsasecurity.com

13 TB RAID Enclosure

Winchester Systems has unveiled the FlashDisk OpenRAID system, which uses new 181 GB drives, for up to 1.45 TB in a single enclosure. Coupled with the 82" high FlashDisk Data Center Cabinet, which holds nine rackmount units and a UPS, FlashDisk OpenRAID can be configured to store over 13 TB in under 6' of floor space, the company says.

Performance for the FlashDisk OpenRAID is clocked at 13,000 I/Os per second and up to 121 MB/s sustained data throughput using two available host ports. The high random I/O makes FlashDisk OpenRAID appropriate for applications such as OLTP, database, manufacturing, and any other requiring elevated levels of random access, Winchester says.

FlashDisk OpenRAID is a "true

hardware, plug-and-play storage solution" the company claims, and does not rely on proprietary host software, RAID managers, software drivers or agents of any kind. It supports RAID 1,3,5, 1+0 and RAID 15. The unit comes with standard Fibre Channel or SCSI interfaces, and connects to host servers from Sun, Hewlett-Packard, IBM, and Intel-based servers.

Pricing for a single 8-drive FlashDisk OpenRAID unit starts at $53,943, or under $0.04 per MB, including cables and a UPS.

Winchester Systems Inc.
149 Middlesex Turnpike
Burlington, MA 01803
http://www.wynsys.com

CD/DVD Library Supports DVD-R/DVD-RAM

Plasmon has announced that its D-Series libraries can now incorporate both DVD-Recorders and DVD-RAM drives in a single unit using a new media rotation mechanism that effectively doubles the capacity of the media.

The new Plasmon robotic media Rotation Station mechanism allows access of double-sided media and increases the 4.7 GB DVD-R and DVD-RAM media to 9.4 GB per disc. Together, the drives and robotics combine to extend D-Series capacity to more than 4.1 TB.

In addition, new 4.7 GB DVD-RAM drives are now fully integrated with the 120 mm format D-Series libraries, almost doubling previous jukebox capacities. The new generation drives also feature improved data transfer rates and access times. In addition, mixing DVD-RAM drives with DVD-R in the same jukebox provides customers with increased flexibility, says the company.

All drives are fully SCSI-enabled, allowing multiple concurrent operations without the limitations of using traditional ATAPI drives in an automated environment. Coupled with the "designed-in" SCSI robotics control system, customers receive greater performance and the ability to communicate with the drives across the SCSI bus through the jukebox control system, Plasmon says.

The new D-Series libraries are priced starting at $14,995. This entry-level configuration includes one DVD-Recorder, one DVD-RAM drive and 120 mediaslots. A high-end, six-drive DVD-R unit with 480 slots is priced at $54,995, and includes the SCSI interface, disc magazines and disc import/export mail slot, the company says.

Plasmon plc
9625 West 76th St.
Eden Prairie, MN 55344
http://www.plasmon.com

RAID Enclosure Uses ATA Drives

Promise Technology has released the UltraTrak100 TX4 and TX8 RAID storage subsystems, which are designed to work with standard off-the-shelf Ultra ATA/100 disk drives rather than more costly SCSI drives. Promise predicts that its RAID enclosures can be configured for almost a third of the price of a comparable SCSI-based system.

The 4- and 8-drive enclosures work with any standard Ultra2 SCSI LVD controller, but are populated with either ATA 100 or ATA/66 drives, which appear as standard SCSI drives. The systems can be configured for RAID 5, 3, 0+1, 1, and 0 configurations.

The UltraTrak TX series includes an internal RAID controller mounted directly into the enclosure. The controller is based on an Intel I960 RISC microprocessor, and includes 16 MB cache memory. Audio alarms, LEDS, and an LCD panel are used to report on the status of drives, arrays, and the enclosure. A management utility offers event logging, disk synchronization, array rebuild and real-time usage statistics. Drives slide easily in and out of the enclosure in the event of a failure, while redundant cooling fans and hot-swappable power supplies prevent overheating and power outages.

The UltraTrak TX series is operating system and platform-independent, working with any system that has embedded or add-in SCSI controller support. List price for the UltraTrak100 TX4 and the TX8 is $1,495 and $2,495.

Promise Technology Inc.
1460 Knoll Circle
San Jose, CA 95112
http://www.promise.com

Edge Router Support IPv6

Internet connectivity provider InterNetShare has announced the availability of its All Aboard! Advanced Edge Router (AER) family of secure, multimedia edge-routing devices for both the classic IPv4 Internet, as well as IPv6.

Once installed, users never have to worry about transitioning network devices to IPv6, the company explains, as AER products automatically detect IPv6-based LAN devices, and provide auto-configuration and transparent connectivity. This lets users preserve their investment in IPv4 devices, while adding IPv6-based systems to the same network segment.

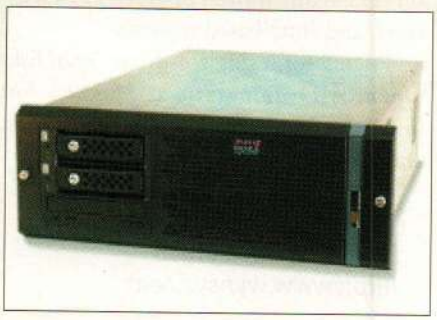
Specifically, AER products can be configured to deliver: NAT-PT (Network Address Translation-Protocol Translation), which converts private IPv4 packets into public IPv6 packets; SIIT (Stateless IP/ICMP Translation Algorithm); DNS-ALG (DNS extensions to Network Address Translators); 6-to-4 (connection of IPv6 Domains via IPv4 Clouds); IPsec support for both Shared Secret and Public Key mechanisms; packet filtering (firewall) functionality on a per protocol, per route basis; a native IPv6 router; HTTP-based configuration and management system; and H.323 routing services, according to the company.

The product family can be distributed either as a complete device for ISPs, or as portable source code for OEMs. Source code is available for Linux, embedded Linux or other popular embedded operating systems. A packaged solution is priced at $4,995 per unit, while the Source code is licensed at $75,000 royalty-free.

InterNetShare Inc.
505 West Olive Ave., Suite 321
Sunnyvale, CA 94086
http://www.internetshare.com

Rackmount UltraSPARC-III Server

Rave Computer Association has announced its new UltraSPARC-III rackmount server, available in single- or dual-processor configurations, and based on the UltraSPARC-III 750 MHz chip. Called the Redundant Rackmount-4UAX2200, this server is designed for applications that require



highly reliable compute-intensive performance, the company says.

Base units come standard with a Sun Netra motherboard, single or dual processors with 8 MB cache, and memory expandable to 8 GB, 1 GB standard. The system can be configured with 4 PCI slots. Serial, USB, Firewire, Fibre Channel and parallel ports are all available, as is on-board 10/100 Mb/s Ethernet, N+1 front-accessible hot swappable power supplies, and three drive bays for CD-ROM, DVD, or other media. The system also comes with four front-accessible 3.5" hard drive bays, and Solaris 8.

The Rave Systems Redundant Rackmount-4UAX2200 boasts a SBus interconnect speed of 4.8 GB/s. With 8 GB of memory installed, the system is ideal for demanding technical applications including data analysis, digital imaging, Internet and telecommunications, the company suggests.

Suggested list pricing starts at $21,750 for a standard configuration of dual 750 MHz processors, 1 GB of memory, and one 36 GB FC-AL disk drive.

Rave Computer Association Inc.
36960 Metro Court
Sterling Heights, MI 48312
http://www.rave.com

Solaris Performance Management Tool

Version 4.0 of Aurora Software's SarCheck performance management tool for SPARC Solaris systems has been announced.

SarCheck identifies performance bottlenecks, finds runaway processes and memory leaks, recommends changes to various tunable parameters and hardware configurations, and quantifies remaining system capacity, using standard operating system tools like `sar` and `ps`. Results of the monitoring utilities are formatted in a plain English or HTML-formatted report.

Version 4.0 lets users graph data collected by `sar` with a third-party graphing tool and control many of the thresholds used by SarCheck's rule set, allowing the system to accommodate a greater number of "unusual conditions," the company says.

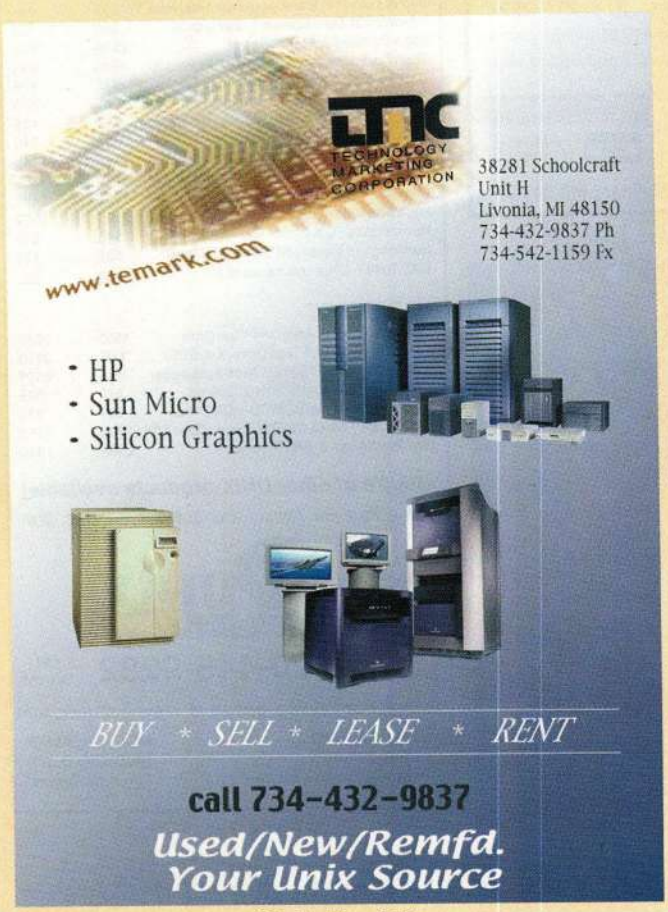
SarCheck runs on many popular UNIX platforms, and support for other languages (Spanish) is planned. Pricing is $625 per system.

Aurora Software, Inc.
P.O. Box 1033
Plaistow, NH 03865
http://www.sarcheck.com

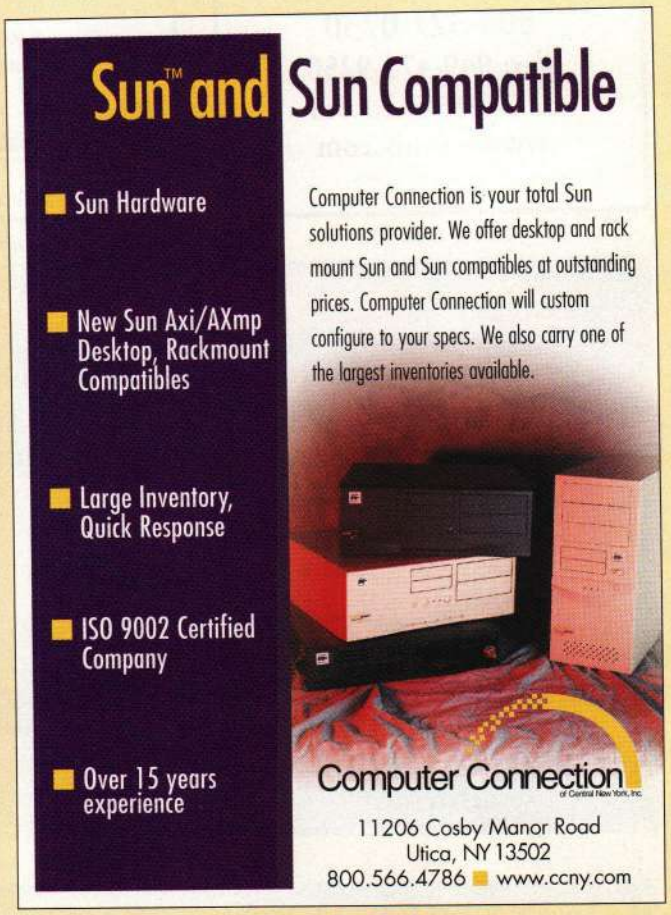
Server/Workstation Marketplace

The high-impact, cost-effective and powerful advertising section reaching over 100,000 technical product specifiers and buyers.

To advertise call Carol Flanagan at **(508) 839-4016** *Email: caflanag@aol.com*



Circle No. 100


Sun™ and Sun Compatible
Sun Hardware
New Sun Axi/AXmp Desktop, Rackmount Compatibles
Large Inventory, Quick Response
ISO 9002 Certified Company
Over 15 years experience
Computer Connection is your total Sun solutions provider. We offer desktop and rack mount Sun and Sun compatibles at outstanding prices. Computer Connection will custom configure to your specs. We also carry one of the largest inventories available.
Computer Connection
11206 Cosby Manor Road
Utica, NY 13502
800.566.4786 www.ccny.com

Circle No. 101

SUN		HP		IBM	
420 System, UNused	$11,350	N4000, New	$59,900	RS6000 Model S80, UNused	$29,500
450MHz, 512MB, 18GB, DVD		2) 550MHz, 1GB, 9GB drive		4-way, 512MB, 9GB disk	
X7005A, UNused	$1,195	L2000, New	$38,400	4193 memory, 8GB, UNused	$36,900
X5242A, Unused	$1,095	2) 440MHz, 1GB, 9GB drive		4191 memory, 2GB, UNused	$13,000
Ultra 10 System, used	$1,700	L1000, New	$12,950	7133-D40 w/16 9GB drives, used	$23,900
300Mhz, 128MB, 4GB, CDROM		440MHz, 512MB, 9GB drive		3570-B1A, used	$2,595

TriniComp SYSTEMS

2 Centennial Drive, Peabody, MA 01960

fax: 978-531-6037
email: sales@trinicomp.com
Call 978-531-7366
www.trinicomp.com

All trademarks are property of their respective companies
Product subject to availability. Pricing subject to change without notice

Circle No. 102

EXPRESS
Computer Systems

BUY – SELL – RENT
SUN MICROSYSTEMS
IBM RS/6000

- SYSTEMS
- PERIPHERALS
- DISK UPGRADES
- PARTS
- MEMORY UPGRADES
- PROCESSOR UPGR.

800-327-0730
Fax: 949-474-9250
sales@ecsunix.com
www.ecsunix.com

Circle No. 103

Buy/Sell new, used and refurbished SUN systems & components

Pyramid Technology

Sun microsystems

E450 2x300MHz, 1GB	$7,950
Ultra 1, 2, 5, 10, 30	$CALL
X7005A 512 MB memory	$995
X7026 2GB memory	$CALL
X7004 256MB UNused	$695

Call for a custom quote.

Toll Free: 1-877-BUY-PTSI (289-7874)
Fax: 978-897-9611
E-Mail: sales@pyramiddec.com
Web site: www.pyramiddec.com

PYRAMID Technology Services, Inc.

EXPORTS a specialty

10 Riverbank Road, Maynard MA 01754

Circle No. 104

SUN - IBM - HP

Hardware & Software for any UNIX Operating System

SOFTWARE	LIST	OURS
WordPerfect 8.0 1st user	495	395
Corel Draw 3.5 1st User	595	495
VSI*FAX 4.0 - Fax Server w/ 5 clients	3235	2750
VSI*FAX 4.0 - 25 Additional Clients	1250	995
FrameMaker 6.0 1-user Personal License	1695	1495
FrameMaker 6.0 1-user Shared License	2495	2275
Adobe Acrobat Distiller Server 100-user	5500	4950
Informix On-Line Dynamic Server 1-usr (min. 10)	1500	call
Informix On-Line WorkGroup Serv 1-usr (min. 5)	295	call
Informix 4GL Development 1-user (min 3)	900	call
Lotus Domino Application Server	2200	1925
Term Communications Software / file transfer	695	595
LP Plus Print Manager/Spooler 4-printer	975	875
FacetTerm Character Windowing S/W 16-user	995	860
Spatch paging SW 4-user/120 subscribers	499	425
DoubleVision Remote Support S/W	895	740
Gear Multimedia CD-RW Recording S/W	1495	1145
Microlite BackupEdge backup/compression S/W	400	350
Cheyenne Arcserve/Open Backup S/W	1995	1495
Legato Networker Backup S/W	2500	1950
Hummingbird PC X-Server Win NT/95	545	395
Hummingbird Maestro NFS Server - Win NT	535	420
MKS Toolkit- UNIX commands for NT	400	335

HARDWARE		
HP 24GB External 4mm DAT Tape Drive	1550	1020
Tandberg 70GB DLT Tape Drive, Ext. SCSI	3995	3450
Exabyte Ix17 420-1000GB Tape Autoloader	7995	6595
IBM 18.3GB External SCSI HD 7200rpm	675	525
Plextor 40X External SCSI CD-ROM Drive	299	220
Plextor 32X/12X/4X CD-RW, Ext. w/ Gear S/W	1995	1550
Digi PortServer II ENET 16-ports	2395	1890

1000's of other UNIX products available!

Other Platforms: LINUX, SCO, SUN X86, NT, DEC, SGI

GO2UNIX.COM™

Free Shipping! 1-800-445-9292 Fast Delivery!

Search our catalog online at www.go2unix.com

Tel: 201-847-8880
Fax: 201-847-8899
E-mail: sales@go2unix.com

MC, VISA, AMEX, DISC, C.O.D. or Purchase Orders

Circle No. 105

Information Requests

For more information on products and services advertised in this section, please circle the appropriate RS number on the Reader Inquiry card located in this issue.

OR

E-mail requests to: ica@ica.cc
Include:
1. Magazine name and issue
2. Your name and address
3. Reader Inquiry number. For more than one inquiry please separate numbers by commas.

anything & everything 4 Sun

On Line ...
www.anythingSun.com
410-820-8100
Ames Sciences, Inc.
501 South Street, Easton, MD 21601
Fax 410-820-8179
sales@anythingSun.com

Circle No. 106

www.workstationcables.com

- Fiber Optics
- IEEE 1394 Firewire
- Keyboard Cables
- Networking Cables
- Serial Cables
- SCSI Cables & Adapters
- USB Products
- Video Cables & Adapters

Sun Developer

Your #1 site devoted to workstation and Sun® cables, adapters and terminators! So check us out today at:

www.workstationcables.com

workstationcables.com
A Division of Lastar Datacomm Solutions
1501 Webster Street
Dayton, Ohio 45404

Phone: 800-293-4971
Local: 937-223-6383
24-Hour Fax: 800-331-2841
Mon. - Fri. • 8:30 am to 6:00 pm EST

Circle No. 107

Comtek Computer Systems

Specializing in :

SUN

SUN Microsystems & Peripherals
In-house OEM Depot Repair
Same Day Advance Exchanges
Spare Parts & Systems Sales
System Lease and Upgrade Programs
Inventory Mgt Programs

- Same day shipping
- One year warranty
- In house technical support
- Authorized Software Support Provider Agent
- On hand inventory
- Customized supply solutions to meet your needs
- **We can handle all your SUN needs**

Comtek Computer Systems

"The first choice in SUN Support"
Give us a call so we can put our highly trained staff to work for you.

1-800-823-4450
WWW.Comtekcomsys.com

Circle No. 108

EVOLVING SOLUTIONS
800.294.4362
Buy Sell Rent

IBM RS/6000
ibmsales@evolvingsol.com
- In House Technical Services
- AIX / ADSM / HACMP Certified
- Upgrades, Parts Features, Short Term Rental

HP 9000
hpsales@evolvingsol.com
- SAN Specialist • HP9000 Certified Tech & Applications
- Custom Enterprise Server • Workstation Configurations
- Engineer On Staff • Disaster Recovery Services On-site
- ISP / Web Host Solutions Provider

SUN MICROSYSTEMS
sunsales@evolvingsol.com
- Custom Enterprise Servers & Workstation Configurations
- Sun Certified Tech & Application Engineer on Staff
- Disaster Recovery Services On-site
- ISP / Web Host Solutions Provider
- Sun Hardware & Applications Training • SAN Specialist

Main: 763.516.6500 • Fax: 763.516.6555 • info@evolvingsol.com

Circle No. 109

RS/6000 **Sun Microsystems**

AS/400 **HP9000**

www.flagshiptech.com

FLAGSHIP TECHNOLOGIES INC.

BUY, SELL, RENT, LEASE

Flagship Technologies is your best source for AS/400, RS/6000, HP9000 and Sun Microsystems equipment and service. We staff certified technicians, supply new and used parts, offer OEM parts, and conduct advance exchange as well as complete upgrades.

Minneapolis (AS/400,RS/6000 & Sun Micro). . **800-416-8900** 763-516-1300
Atlanta (HP9000). **888-740-4600** 770-576-1902

Circle No. 110

Get in Touch with Trident.

TOUCH SCREEN TECHNOLOGY EXCELS AT
Simplifying computer input
Saving space
Supporting fast, frequent computer operations
Reducing maintenance costs and down-time

AND NOW, YOU CAN HAVE TOUCH WITHOUT MODIFYING YOUR EXISTING APPLICATIONS. TRIDENT'S TOUCH SOLUTIONS INCLUDE
Touch-configured monitors, X terminals and Netstations
Touch screen integration with your monitors
Touch screen hardware and driver software
All major touch technologies
Rack mounting or ruggedized NEMA enclosures and kiosks with touch
Touch application software development services

TRIDENT'S TOUCH SOLUTIONS WORK WITH
UNIX Workstations X Terminals UNIX PC's
DOS & Windows Netstations Macintosh

Trident puts you in touch.™

TRIDENT TECHNOLOGY SOLUTIONS

TRIDENT SYSTEMS INC. email: touch@tridsys.com
TEL 703.273.1012 url:http://www.tridsys.com/cpg
FAX 703.273.3763

Circle No. 111

GULFCOAST

Ask about the new line of midframe Sun Fire™ servers.

We provide high-performance Sun™ servers. And the support team to go with them.

At Gulfcoast, we offer more than an extensive inventory of Sun™ servers, Sun storage solutions, and parts. You can also count on us for outstanding customer service, including same-day quotes, quick delivery, and ongoing Help Desk support. And our staff of certified engineers will help integrate, configure and optimize a solution to fit your environment.

For more information, contact Gulfcoast today at 1-800-587-7882, sales@gcw.com, or www.gcw.com

Sun, Sun Microsystems, the Sun Logo and Sun Fire are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and other countries.

Circle No. 112

Ph# 800-426-3090
Fx# 847-317-2622
www.meridianleasing.com

MERIDIAN
TECHNOLOGY LEASING SERVICES

570 Lake Cook Road
Suite 300
Deerfield, IL 60015

Product Manager:	Al Cosenza	acosenza@meridianleasing.net	847-317-2270
Compaq Servers: DEC Alphaservers: HP:	1500s, 2500s, 4500s, 5000s, 6500s and PC equipment 8400s, 4100s, 2100s, 1000, 1000A, and Storage Works All K Products, Workstations, Disk and Tape		
Product Manager:	Jason Davis	jdavis@n2nsolutions.com	847-592-3969
Cisco:	N2N Solutions, acquired by the Meridian Group in July, 1999, designs and supports end-to-end internetworking solutions, from systems analysis to support. A high-profile business partner of Cisco Systems, Inc., N2N Solutions is authorized to sell Cisco products and services throughout the United States.		
Product Manager :	Tom Haug	thaug@meridianleasing.net	847-317-2198
IBM RS/6000: Sun Microsystems:	Servers, Workstations, Parts, and Peripherals Ultra Workstations, Enterprise Servers, Parts, and Peripherals		

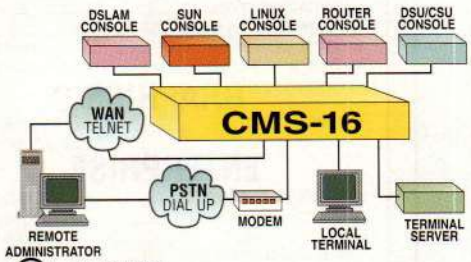
We are now entering our 21st year of meeting the (IT) needs of Corporate America. From leading-edge connectivity to commonplace data center solutions, The Meridian Group is a mainstay in the competitive (IT) leasing and reseller services business. Solidly positioned to take advantage of market changes and opportunity we continue to earn our client's business by consistently delivery reliable service and leading edge technology at reasonable prices.

Circle No. 113

Console Management Over Telnet!

Access Network Serial Console Ports . . . From Anywhere!

The CMS-16 Console Management Switch provides secure, in-band and/or out-of-band access to RS232 console ports and maintenance ports on UNIX servers, routers, and other network equipment. System administrators can access remote devices in order to change configuration parameters, connect users to restricted ports, collect buffered data, and perform a variety of other control related functions.



- ✔ Sixteen (16) RS232 DB9 Serial Ports
- ✔ 10Base-T Ethernet Port
- ✔ TCP/IP Security Features
- ✔ Port-Specific Password Protection
- ✔ Non-Connect Port Buffering
- ✔ Modem Auto-Setup Command Strings
- ✔ Co-location Features
- ✔ AC and 48V DC Powered Models

Visit website for complete NetReach™ product line.

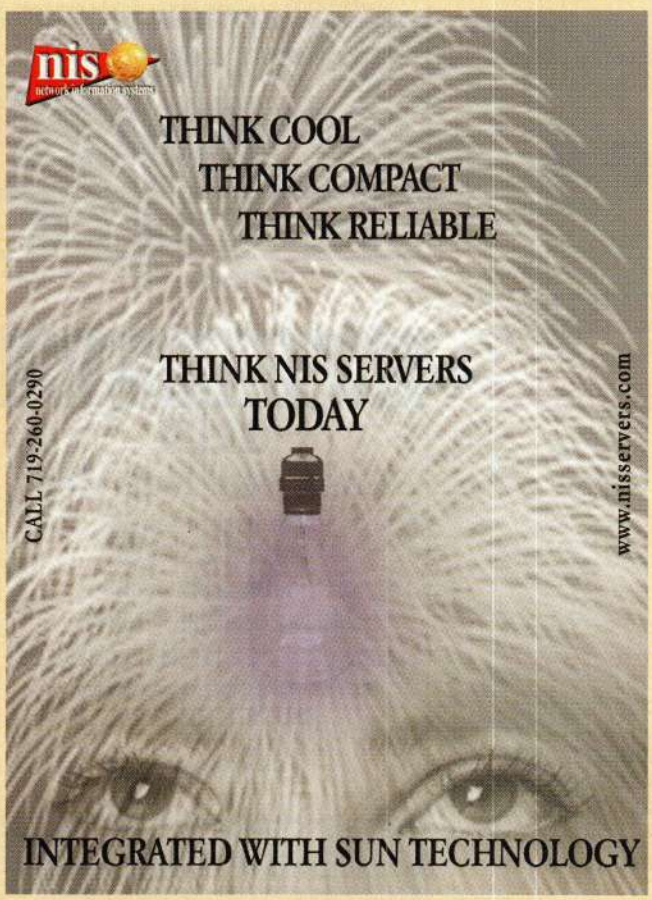
c UL us LISTED CE NEBS-3

wti western telematic inc.

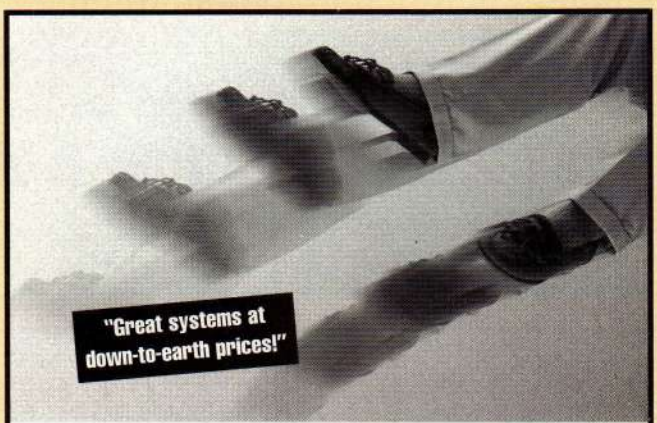
(800) 854-7226 • www.wti.com

5 Sterling • Irvine • California 92618-2517 • (949) 586-9950 • Fax: (949) 583-9514

Circle No. 114



Circle No. 115



Customer Blown Away By Low Prices

Work Stations • Servers • Peripherals • Parts
Sales • Rentals • Leases • Exchanges • Repairs
Maintenance • 120 Day Warranty

Call today to buy, sell or trade SUN, HP and Silicon Graphics equipment with Security!

SECURITY COMPUTER SALES

2340 County Road J. • White Bear Lake, MN 55110
651/653-5200 • 651/653-4300 • Sales@Securitycomputer.com

Circle No. 116

Sun

BUY SELL TRADE

Workstations and Servers

Boards Monitors Disks Tapes Mass Storage Solutions

♦ Same Day Shipping ♦ Competitive Pricing
♦ Custom Configurations Available

DEPOT REPAIR CENTER

♦ 30% To 50% Cost Savings ♦ One Year Warranty
♦ Overnight "Swap" Service

ACC, Inc.
AdvanTec Computer Company
48989 Milmont Drive Fremont, CA 94538
Phone: (510) 687-9100 Fax: (510) 687-9733
E-mail: sales@advanteco.com
http://www.advanteco.com

Circle No. 117

HIGH IMPACT! COST-EFFECTIVE!
Advertise in
MARKETPLACE
CALL CAROL FLANAGAN AT
(508) 839-4016

SYSTEMS & CPU's

Ultra 1/200E-C3D-128mb-2.1GB-CD-FL	**$1595.00**
Ultra 2 dual 200mhz-C3D-128mb-4.2GB-CD-FL	$2495.00
Ultra 2 dual 300mhz-C3D-128mb-4.2GB-CD-FL	$3395.00
Ultra 60-dual 300mhz-C3D-128mb-18.1GB-20"	$3895.00
Ultra 60-dual 360mhz-C3D-128mb-18.1GB-20"	$5295.00
Enterprise 450 Dual 300mhz, 512mb,18GB	**$8415.00**
Enterprise 250 Dual 300mhz, 512mb, 72GB	**$7295.00**
Enterprise 3500, 2 X400mhz\4mb, 512mb, 18GB	**$28,900.00**
Ultra 5/400mhz, 20GB disk, 128mb ram, PGX, CD	$2100.00
S5/170-32mb-1.05GB-TX1N 170mhz	$795.00
X1191A 300mhz Ultrasparc cpu	$995.00
X1192A 360mhz Ultrasparc cpu	$1595.00
X1195A 450mhz Ultrasparc cpu	$3295.00
X1194A 400mhz\2mb cache Ultrasparc cpu	$1895.00

MEMORY & DISKS

Sun 18.1GB U60\E250\MP\Array	$375.00
Sun 2GB SS5\20\Ultra	$165.00
Sun 4GB SS5\20\Ultra	$250.00
Sun 9GB Ultra\Array	$295.00
Sun X7004A 256mb kit\Ultra	$825.00
Sun X7022A 256mb EX000\EX500	$595.00
Sun X7023A 1GB memory kit EX000\EX500	$3895.00
Sun X7005A 512 memory for U80\E450	**$2295.00**
Sun Storage Arrays 114/RSM/A5000/A5100/A5200	Call

GRAPHICS

TX1/TX4/**MGX+ 4mb 24 bit**	$125/$595/$175
Creator/Creator3D sbus/C3D PCI	$165/$195/$475

ENTERPRISE

EX000 X2601A system board	$2695.00
EX000 X2602A system board	$3695.00
EX000 X2611A I/O board	$2695.00
EX000 X2560A 336mhz\4mb cpu	$2295.00
EX000 X2550A 250mhz\4mb cpu	$2295.00
EX000 X2570 400mhz\4mb cpu	**$2895.00**
EX000 X2580A 400mhz/8mb cpu	**$6995.00**

Sun Ultra 10\333-256mb-9.1GB-CD-FL-PGX-20" color $1975.00
Sun X7124A 24" color monitor $1595.00

Large inventory of hardware including spare parts, servers, workstations, peripherals, memory, and 3rd party products
NEW! industry leading 24 month warranty!!!
Check our web page for weekly specials -
www.gshiis.com Email: sales@gshiis.com
*** Call for full product line and availability. Same day shipping on all stock parts. Trades accepted. Full line of leasing and short term rental options available.**

VISA American Express Card MasterCard SPARC International Member

GSH Intelligent Integrated Systems, Inc.
95 Fairmount Street Fitchburg, MA 01420
Tel: **(603) 529-7880** Fax: **(603) 529-7884**

Circle No. 118

LOWEST PRICES ON SUN MEMORY

Clearpoint Enterprises, Inc.

Save time, money dealing directly with Manufacturer

"If My Memory Serves Me Right ...It Must Be Clearpoint"

34 Hayden Rowe St.
Hopkinton, MA 01748
(800) 253-2778
(508) 544-9694
Fax (508) 544-8206
www.clearpoint.com
Email: memory@clearpoint.com

Circle No. 119

WANTED TO BUY

SUN

Cash Paid Now For
All Enterprise Servers
Ultra 1 & 2, 30 & 60
SUN Disk & Disk Arrays
Memory & Enterprise options

SAVE on OFF LEASE & REFURBISHED SYSTEMS
Systems available for sale, lease or rent
CALL NOW FOR QUOTE

DATALEASE

(800) 456-6233
FAX (714) 632-9248
sales@datalease.com

Circle No. 120

mini·computer exchange
www.mce.com

We do SUN SGI CISCO

BUY IT
SELL IT
TRADE IT
RENT IT
REPAIR IT
SERVICE IT
REFURBISH IT
WARRANTY IT
FINANCE IT

Since 1973

Servers, Workstations & Peripherals

Custom Systems & Components

888-733-4400
408-733-4400
FAX 408-733-8009
610 N Pastoria Ave.
Sunnyvale, CA
94086 USA

MCE WORLD MARKETPLACE

Circle No. 121

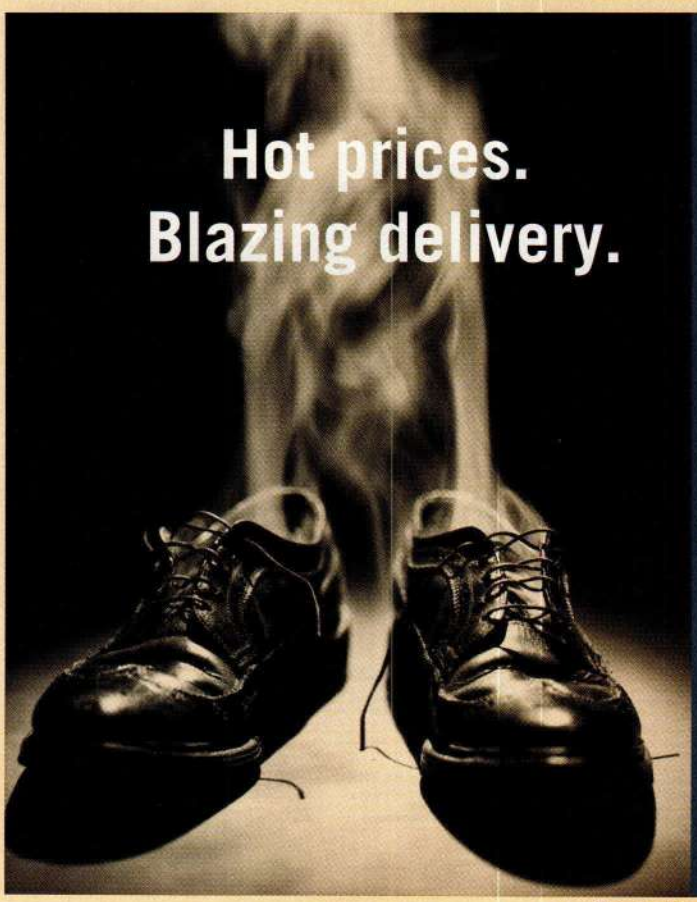
Customer Goes Ape Over Great Service

Call today to buy, sell or trade SUN, HP and Silicon Graphics equipment with Security! Unbeatable service, unbeatable value -- No wonder so many people are going ape.

SECURITY COMPUTER SALES

2340 County Road J. • White Bear Lake, MN 55110
651/653-5200 • 651/653-4300 • Sales@Securitycomputer.com

Circle No. 122



Save up to 60% or more on used Sun Microsystems equipment.

Solar Systems delivers red-hot selection, prices and expertise you just won't find anywhere else. With our huge inventory of pre-tested equipment, we can quickly ship custom configurations that solve your problem the first time around. Whether you need a performance upgrade, replacement parts or a complete system, we turn up the heat to get it to you fast. And your satisfaction is guaranteed.

Buying or selling Sun Microsystems equipment? Call for a free quote today at (800) 253-5764, ext. 22 or visit www.solarsystems.com/22.

solarsystems

Circle No. 123

DATAseek

Making Information Available

Formerly Encore Technology Group

Sales/Rental/Leasing
- New & Recertified RS/6000
- In-house Technical Support
- Short-term Rental & Leasing
- Same Day Shipping

Sun IBM HP

Now Specializing in Storage Solutions
- Storage Network Design
- NAS Installation
- Clustering Solutions
- SAN Consulting & Installation

VERITAS Exabyte BROCADE

Our Online Catalog is Now Available at:
www.dataseekonline.com

email: sales@dataseekonline.com

800.880.0202

YOUR COMPLETE SATISFACTION IS THE MEASURE OF OUR SUCCESS.

Circle No. 124

EXA- Bitten ??
DLT- Bitten !!

Rent or Buy refurbed,or
Send Your Repairs to the
BEST Quality Source for
8mm & DLT Tape Drives and Libraries

SUN Microsystems Servers
Rent, Buy, Sell, Spare Parts, Repairs
IMMEDIATE DELIVERY

WEX

West Coast Computer Exchange, Inc.
11167-A Trade Center Drive
Rancho Cordova, CA 95670

Call: 916-635-9340 or FAX: 916-635-9485
Catalog at Web Site: www.wccx.com
email: sales@ wccx.com

Circle No. 125

GENUINE SUN REMANUFACTURED EQUIPMENT

PROVEN PRODUCTS FOR YOUR BUSINESS

Sun Genuine Remanufactured Product

FOR SUN CUSTOMERS LOOKING FOR HIGH VALUE, HIGH PERFORMANCE ALTERNATIVES TO NEW EQUIPMENT, THERE IS ONLY ONE SOURCE FOR WARRANTY BACKED, SUN ENGINEERED SOLUTIONS: GENUINE SUN REMANUFACTURED EQUIPMENT THAT BEARS THIS LOGO*.

SAMPLE SELECTIONS: (*MULTIPLE QUANTITIES AVAIL.*)

ULTRA 2: 2X300MHZ, 256MB,9.1GB,32XCD	$3,500
ULTRA 60: 2X300MHZ, CR3D3, 2X9GB 7200RPM, 512MB	$3,800
ENTERPRISE 3500 SERVER W/4X400MHZ, 2.5GB MEMORY	$67,800
SUN STOREDGE A1000 TABLETOP 12X9.1GB 10KRPM	$11,900
ULTRA 10 300MHZ, 4.3GB, 24XCD, 128MB	$1,500

LOG-ON TO WWW.SUN.COM/IBB/REMANUFACTURED/ AND FIND A HIGH VALUE, HIGH PERFORMANCE SOLUTION THAT FITS YOUR NEEDS TODAY.

Sun microsystems

take it to the n[th]

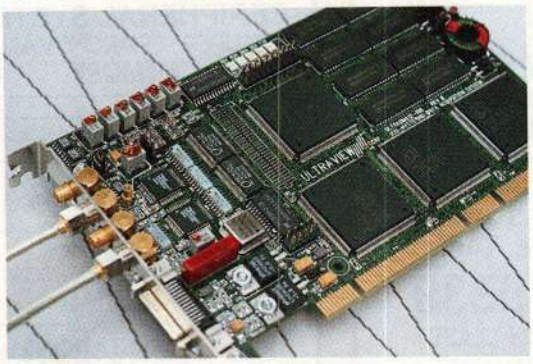
*THE GENUINE SUN REMANUFACTURED LOGO ENSURES THAT EVERY MAJOR COMPONENT WE SHIP MEETS THE CURRENT PRODUCTION AND TEST SPECIFICATIONS FOR NEW EQUIPMENT. ALL GENUINE SUN REMANUFACTURED IS SUPPORTED BY A FULL SUN WARRANTY THAT PROTECTS YOUR INVESTMENT FOR YEARS TO COME.

©2000 Sun Microsystems,Inc.All rights reserved.Sun,Sun Microsystems,the Sun logo, Solaris, Ultra, and We're the dot in .com are trademarks or registered trademarks of Sun Microsystems,Inc.,in the United States and other countries.

Circle No. 126

FAST PCI DMA A/D

64 Bit DMA



100 MS/s DATA ACQUISITION

Acquire gigabytes of continuous analog data directly to Sun memory at blazing speed.

- AD12-100DMA - Dual 12-bit 25 MS/s A/D
- ADDA12- 100DMA - Dual 12-bit 25 MS/s A/D + Dual 14-bit 15 MS/s D/A
- AD8-100DMA - Dual 8-bit 50 MS/s A/D (or single 8-bit 100MS/s A/D)

ULTRAVIEW CORP.
34 Canyon View
Orinda, CA 94563
Tel: (925) 253-2960

www.ultraviewcorp.com
sales@ultraviewcorp.com
Fax (925) 253-4894

Circle No. 127

Customer's Head Spins Over Great Deals

Call today to buy, sell or trade SUN, HP and Silicon Graphics equipment with Security! Unbeatable service, unbeatable value -- No wonder so many people are going ape.

SECURITY COMPUTER SALES

2340 County Road J. • White Bear Lake, MN 55110
651/653-5200 • 651/653-4300 • Sales@Securitycomputer.com

Circle No.128

Investment Counseling: Sun Technology Ownership.

Appreciate Depreciation

G–2 Systems sells and buys quality refurbished Sun Microsystems equipment.

Innovative solutions to your Sun Microsystems technology lifecycle challenges.

G-2 Systems

G–2 Systems
55 Chase Street, Methuen, MA 01844
P: 978.681.6010 F: 978.681.4650
E: sales@g-2systems.com www.g-2systems.com

Circle No. 129

RENT to OWN

SUN(FIRE) 280R Server $2,400/mo*

A35-WRF2-8GQB1
- 2 X 750Mhz processors
- 8Gb Memory
- 2 x 36.4Gb 10K FCAL Hard Drive
- 1 X DVD
- 2 X Power supplies

Sun list price $71,995.00

*Lease for 36 months with $1.00 buyout.

Shorter term rentals (3,6,12 mo) available
CISCO Routers and Hubs available
Subject to supply on hand and credit approval

DATALEASE

(800) 456-6233
FAX (714) 632-9248
sales@datalease.com

Circle No. 130

Information Requests

For more information on products and services advertised in this section, please circle the appropriate RS number on the Reader Inquiry card located in this issue.

OR **E-mail requests to: ica@ica.cc**
Include:
1. **Magazine name and issue**
2. **Your name and address**
3. **Reader Inquiry number. For more than one inquiry please separate numbers by commas.**

RENT TO OWN

SUN E420R Server

$795/mo*

A33-SI-ULD2-2GFB1
2 x 450Mhz processors
2Gb Memory
36Gb Internal disk
List price $29,995.00

* Lease for 36 months with $1.00 buyout or purchase for $22,500.00.
CISCO Routers and Hubs available.
*Subject to supply on hand and credit approval.

DATALEASE

(800) 456-6233
FAX (714) 632-9248
sales@datalease.com

Circle No. 131

UNIVERSAL CAPITAL FUNDING

SUN MICROSYSTEMS
WORKSTATIONS
PRINTERS SERVERS MEMORY
OPTIONS DISK TAPE

800-883-9339

BUY • SELL • RENT • LEASE •
UPGRADE • EXCHANGE

UNIVERSAL CAPITAL FUNDING
15695 Medina Road, Plymouth, MN 55447
VOICE (763) 577-4200 FAX (763) 551-9311
EMAIL: lyonts@unicap.com

Circle No. 132

ATLANTIC

PERIPHERAL SYSTEMS
"SUPER" SPARC™ VALUES!

MEMORY
64MB SS20——$250
64MB Ultra 1——$275
128MB Ultra 2——$495
256MB Kit Ent. Server——$1,095

STORAGE
2.1GB SGT 'cuda Int/SCSI——$425
4.3GB SGT 'cuda Int/SCSI——$650
9.0GB SGT 'cuda Int/SCSI—— $950
9.0GB SGT Cheetah Int/SCSI $1,295

SPECIAL!
23GB SGT Elite Int/SCSI——$1,895

WORKSTATIONS
SS Ultra 1, 2——CALL FOR QUOTE!

NOW! VISIT OUR WEBSITE!
WWW.ATLANTICPS.COM
912-929-7099 FAX 912-929-7098
E-MAIL: atlanticps@mindspring.com

Circle No. 133

RS/6000 Products

ExpressPoint
Technology Services

R/S 6000
SAVINGS UP TO 75%
- Whole machines
- Parts
- Repairs
- Logistics management
- Features
- Upgrades

TOLL FREE
877-397-7646

Jeff Karschnik x5760
email: jeffk@expresspoint.com

Circle No. 134

RS/6000
9076/SP2
NEW • REFURBISHED
BUY • SELL • LEASE
SYSTEMS • PARTS
PERIPHERALS

Worldwide Trade Corporation
1363 Park Road
Chanhassen, MN 55317
Call (952) 474-0322
Fax (952) 474-1940
www.RS-6000.com
e-mail: info@rs-6000.com

Circle No. 135

WebServer ONLINE MAGAZINE

The latest information to help Web professionals manage and optimize their site is just a click away.

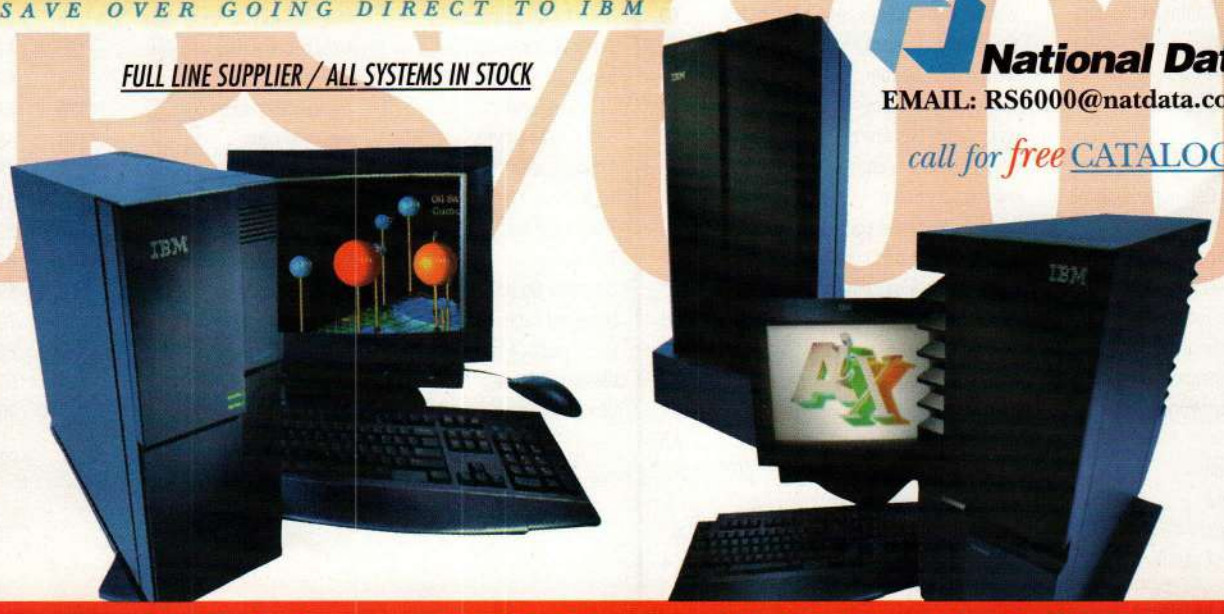
http://webserver.cpg.com/

Why Pay More?

IBM RS/6000

SAVE OVER GOING DIRECT TO IBM

FULL LINE SUPPLIER / ALL SYSTEMS IN STOCK



National Data

EMAIL: RS6000@natdata.com

call for free CATALOG

CALL : 800 - 406 1811

Circle No. 136

CAREER OPPORTUNITIES

Our Readers are IT Professionals

Software Developers, System Analysts, Web Site Developers, Systems Integration Managers, Network Directors and more

TO ADVERTISE YOUR POSITIONS CONTACT

CAROL FLANAGAN AT (508) 839-4016

CAFLANAG@AOL.COM

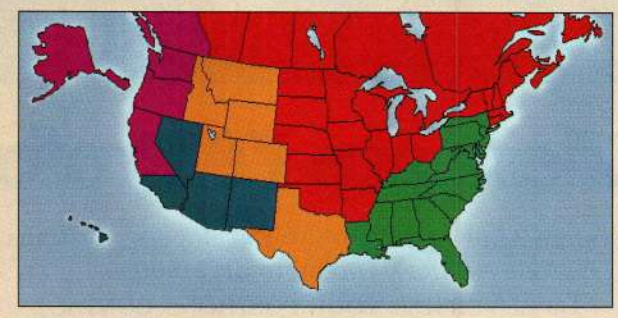
ADVERTISERS' INDEX

Advertiser	Url	Reader Inquiry #	Page
AdvanTec - ACC, Inc.	www.advanteco.com	117	66
Ames Sciences	www.anythingSun.com	106	62
APC	www.apc.com		17
Atlantic Peripherals	www.atlanticps.com	133	70
Clearpoint Enterprises	www.clearpoint.com	119	66
Compaq	www.compaq.com		C4
Computer Connection	www.ccny.com	101	61
Comtek Computer Systems	www.comtekcomsys.com	108	63
Datalease	www.datalease.com	120	66
Datalease	www.datalease.com	130	69
Datalease	www.datalease.com	131	70
Dataseek	www.dataseekonline.com	124	68
Evolving Solutions	www.evolvingsol.com	109	63
Express Computer Systems	www.ecsunix.com	103	62
ExpressPoint	www.expresspoint.com	134	70
Fall Internet World	www.internetworld.com		57
FCTC SAN Conference	www.fctexchange.com		C2
Flagship Technologies	www.flagshiptech.com	110	63
G-2 Systems	www.g-2systems.com	129	69
GSH Systems	www.gshiis.com	118	66
Gulfcoast Workstation	www.gcw.com	112	64
Intel	www.intel.com		4,5
Lastar.com	www.workstationcables.com	107	62
Lightwave	www.lightwavecom.com		1
Linux World	www.linuxworldexpo.com		51
Meridian Leasing	www.meridianleasing.com	113	64
Minicomputer Exchange	www.mce.com	121	67
National Data	www.natdata.com	136	71
Network Information Systems	www.nisservers.com	115	65
Networld & Interop	www.interop.com		C3
Open Systems Express	www.go2unix.com	105	62
Pyramid Technology	www.pyramiddec.com	104	62
Qualstar	www.sansmart.com		27
Rave	www.rave.net		11
Security Computer Sales	www.securitycomputer.com	116	65
Security Computer Sales	www.securitycomputer.com	122	67
Security Computer Sales	www.securitycomputer.com	128	69
Solar Systems	www.solarsystems.com	123	67
Sun Microsystems	www.sun.com	126	68
Tatung Science & Technology	www.tsti.com		15
Technology Marketing	www.temark.com	100	61
Trident Systems	www.tridsys.com	111	63
Trinicomp	www.trinicomp.com	102	61
Ultraview Corp.	www.ultraviewcorp.com	127	69
Universal Capital Funding	www.unicap.com	132	70
West Coast Computer Exchange	www.wccx.com	125	68
Western Telematic	www.wti.com	114	65
Worldwide Trade Corporation	www.RS-6000.com	135	70

For a convenient way to reach all our advertisers, point your browser to http://swexpert.com/market/

The Ad Index is published as a service to our readers.The publisher does not assume any liability for errors or omissions.

ADVERTISING SALES OFFICES



Northern California/ Northwest/Western Canada/Alaska

BILL COOPER

1000 Elwell Court, Ste. 214
Palo Alto, CA 94303
Phone: (650) 961-1760
Fax: (650) 961-4982
email: bcooper@cpg.com

Southern California/ Nevada/Arizona/ New Mexico/Hawaii

GREG CRUSE

3816 Calle Tiburon
San Clemente, CA 92672
Phone: (949) 361-1870
Fax: (949) 361-1367
email: greg@cpg.com

Texas/Colorado/Idaho/ Montana/Wyoming/Utah

LINDA LIEBICH

10931 Sierra Oaks Drive
Austin, TX 78759-5164
Phone: (512) 219-9911
Fax: (512) 219-9988
email: lindal@concentric.net

New England/Upstate New York/Central & Eastern Canada/Midwest

CAROL A. FLANAGAN

80 Worcester St., Ste. 9
North Grafton, MA 01536
Phone: (508) 839-4016
Fax: (508) 839-4226
email: caflanag@aol.com

New York/Southeast/ Mid-Atlantic

JOANNA PARKER

18 Stephanie Drive, Ste. 3
Stirling, NJ 07980
Phone: (908) 542-0789
Fax: (908) 542-0782
email: joanna@cpg.com

The Server/Workstation Marketplace/Classifieds/ Recruitment Ads

CAROL FLANAGAN
Phone: (508) 839-4016
Phone: (508) 839-0720

80 Worcester St., Ste. 9
North Grafton, MA 01536
Fax: (508) 839-4226
email: caflanag@aol.com

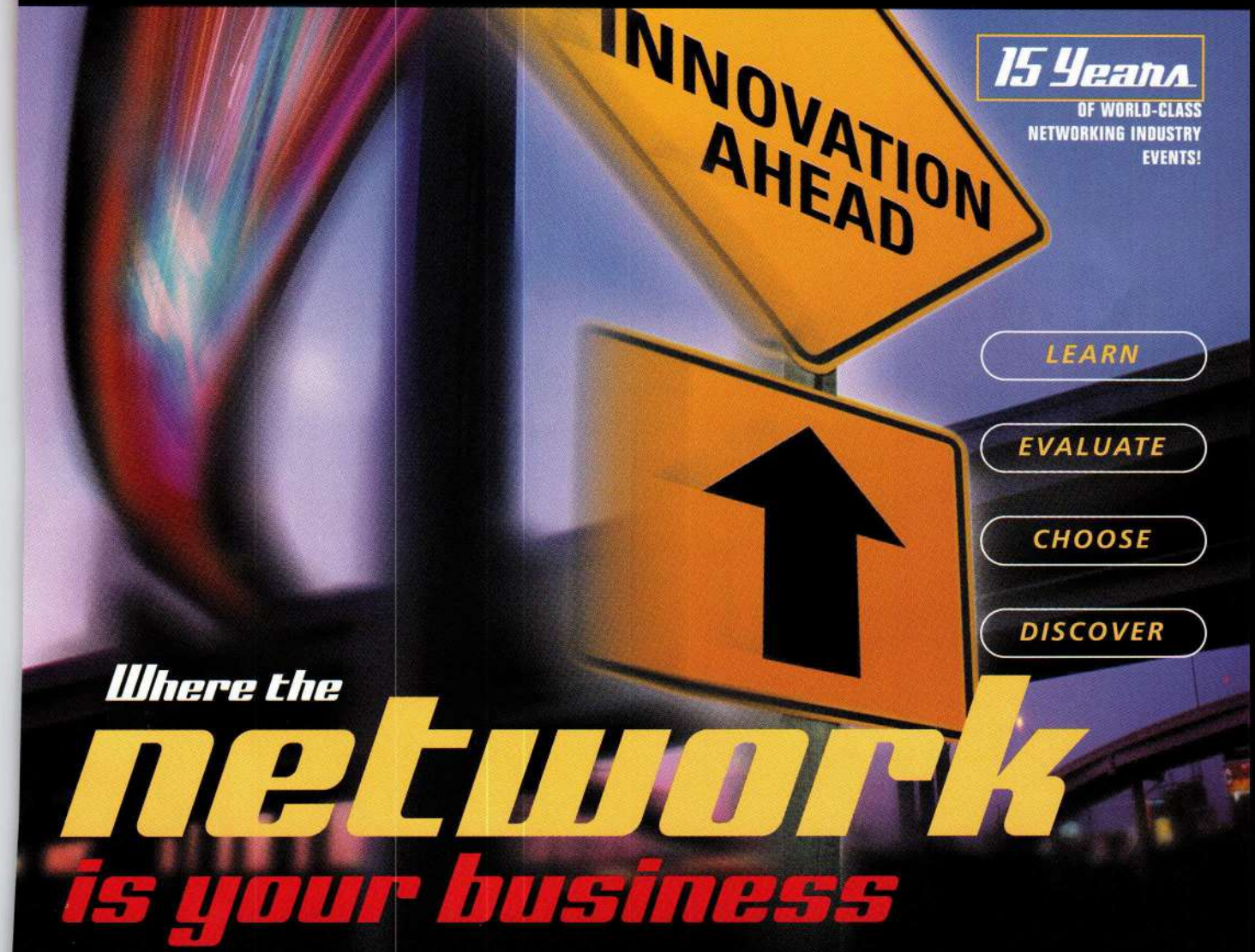
Key3 Media Group

PRESENTS

NETWORLD+INTEROP

Conference: September 9–14, 2001
Exhibition: September 11–13, 2001
Georgia World Congress Center
Atlanta, Georgia

AN INTEROP EVENT



15 Years
OF WORLD-CLASS NETWORKING INDUSTRY EVENTS!

LEARN
EVALUATE
CHOOSE
DISCOVER

Where the network is your business

Get the keys to the latest solutions for driving business success.

- **THE EXHIBITION:** Evaluate network infrastructure and management solutions from 600+ leading vendors.
- **EDUCATION:** Choose from the most current content available in more than 60 conferences, tutorials and workshops.
- **INTEROPNET:** Discover emerging technologies showcased at the InteropNet Labs (iLabs), and cutting-edge solutions working in real time on the InteropNet Event Network (eNet).
- **COMMUNITY:** Connect with other enterprise professionals and service providers and exchange ideas with industry experts.

Register online at www.interop.com

or call 888-886-4057;
international 781-433-1516.

POWERED BY Digital Island.
The Official Internet Provider of Key3Media

Use Priority Code SEMG when registering

OFFICIAL CORPORATE SPONSORS OF KEY3MEDIA GROUP

Blue from American Express, the Official Card of NetWorld+Interop

Mercedes-Benz
Official Automotive Sponsor

NYSE
New York Stock Exchange

Copyright © 2001 Key3Media Events, Inc., 303 Vintage Park Drive, Foster City, CA 94404-1135. All Rights Reserved. NetWorld+Interop 2001 Atlanta is a Key3Media event. Interop is a trademark of Key3Media Events, Inc. NetWorld is a service mark of Novell, Inc., and may be registered in certain jurisdictions.

Compaq | business

100% PURE COMPAQ.
BECAUSE IN THE LONG RUN,

A CHEAP IMITATION IS NEITHER.

$0 down, 0% lease on select Compaq ProLiant Servers before September 30. Why buy a cheap imitation when you can have the real thing? Only Compaq can provide your business with the reliability and service that has made us #1 in customer satisfaction[1]. Designed with your business — and budget — in mind, these ProLiant Servers are long-term investments that pay off right away. And with a great lease offer like this, there's just no reason to settle for second best.

ProLiant DL360 Server

$96 per month[1]

$0 down, 0% lease for up to 30 months[2]

Buy it for $2869*

QuickAccess Code: S1X8X1-17353

Intel® Pentium® III processor 1.00GHz, 256K cache

18GB Pluggable Ultra3 SCSI hard drive[3]

128MB PC 133MHz ECC SDRAM (exp. to 4GB[3])

1U[4] form factor

Dual processor capable

3/3/3 parts, labor, and on-site warranty[5]

- *Network Magazine* Product of the Year 2001

Optional Upgrades: Remote manageability including lights-out capability

ProLiant ML370 Server

$116 per month[1]

$0 down, 0% lease for up to 30 months[2]

Buy it for $3469*

QuickAccess Code: S1X8X1-16654

Intel Pentium III processor 1.13GHz, 512K cache

18GB Pluggable Ultra3 SCSI hard drive[3]

256MB PC 133MHz ECC SDRAM (exp. to 6GB[3])

5U[4] form factor (Tower available)

Dual processor capable

3/3/3 parts, labor, and on-site warranty[5]

Optional Upgrades: On-line spare memory, redundant hot plug fans, hot plug PCI and power supply

ProLiant DL380 Server

$131 per month[1]

$0 down, 0% lease for up to 30 months[2]

Buy it for $3919*

QuickAccess Code: S1X8X1-16754

Intel Pentium III processor 1.13GHz, 512K cache

18GB Pluggable Ultra3 SCSI hard drive[3]

256MB PC 133MHz ECC SDRAM (exp. to 6GB[3])

2U[4] form factor

Dual processor capable

3/3/3 parts, labor, and on-site warranty[5]

Optional Upgrades: On-line spare memory, redundant hot plug fans, hot plug PCI and power supply

intel inside
pentium® III

To buy direct or for your nearest reseller

1-888-349-5721 *toll-free*

compaq.com/purecompaq8

COMPAQ
Inspiration Technology

*Prices shown are Compaq prices; reseller and retailer prices may vary. Prices shown are subject to change and do not include applicable state and local sales tax or shipping to recipient's destination. [1]Based on Technology Business Research, Inc. (TBR) Q1 2001 Corporate IT Buying Behavior and Customer Satisfaction Study: Intel Servers. [2]The monthly lease payments shown are based on the Compaq purchase prices and 30-month lease terms. Monthly lease payment will vary if reseller or retailer prices vary or lessee requests a lease term at less than 30 months. Financing available through Compaq Financial Services Corporation (CFSC) to qualified commercial customers in the U.S. Financing is subject to credit approval and execution of standard CFSC documentation. Other restrictions may apply. CFSC reserves the right to change or cancel this program at any time without notice. These programs cannot be combined with any other special discounting without Compaq's permission. [3]This $0 down, 0% lease offer is a lease with a fair market value purchase option at the end of the lease term. The monthly lease payments are calculated at a 0% implicit lease rate, assuming lessee does not exercise a purchase option at the end of the lease term and timely returns the leased equipment to CFSC at the end of the lease term and disregarding any charges payable by lessee other than rent payments (such as taxes, fees, and shipping charges). [4]For hard drives, GB=billion bytes. [5]1U=1.75-in. [6]Certain restrictions and exclusions apply. Consult the Compaq Support Center for details. Three years on-site and labor not available in all countries. ©2001 Compaq Computer Corporation. All rights reserved. Compaq, the Compaq Logo, and ProLiant are registered trademarks of Compaq Computer Corporation. Inspiration Technology is a trademark of Compaq Information Technologies Group, L.P. in the U.S. and other countries. Intel, the Intel Inside Logo, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. Other products mentioned herein may be trademarks of their respective companies. Compaq is not liable for editorial, pictorial, or typographical errors in this advertisement. 9697A/08/01